(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
9 August 2001 (09.08.2001)

PCT

(10) International Publication Number
WO 01/57182 A2

(51) International Patent Classification[7]: C12N

(21) International Application Number: PCT/US01/01354

(22) International Filing Date: 17 January 2001 (17.01.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/179,065	31 January 2000 (31.01.2000)	US
60/180,628	4 February 2000 (04.02.2000)	US
60/184,664	24 February 2000 (24.02.2000)	US
60/186,350	2 March 2000 (02.03.2000)	US
60/189,874	16 March 2000 (16.03.2000)	US
60/190,076	17 March 2000 (17.03.2000)	US
60/198,123	18 April 2000 (18.04.2000)	US
60/205,515	19 May 2000 (19.05.2000)	US
60/209,467	7 June 2000 (07.06.2000)	US
60/214,886	28 June 2000 (28.06.2000)	US
60/215,135	30 June 2000 (30.06.2000)	US
60/216,647	7 July 2000 (07.07.2000)	US
60/216,880	7 July 2000 (07.07.2000)	US
60/217,487	11 July 2000 (11.07.2000)	US
60/217,496	11 July 2000 (11.07.2000)	US
60/218,290	14 July 2000 (14.07.2000)	US
60/220,963	26 July 2000 (26.07.2000)	US
60/220,964	26 July 2000 (26.07.2000)	US
60/225,757	14 August 2000 (14.08.2000)	US
60/225,270	14 August 2000 (14.08.2000)	US
60/225,447	14 August 2000 (14.08.2000)	US
60/225,266	14 August 2000 (14.08.2000)	US
60/225,213	14 August 2000 (14.08.2000)	US
60/225,759	14 August 2000 (14.08.2000)	US
60/224,519	14 August 2000 (14.08.2000)	US
60/224,518	14 August 2000 (14.08.2000)	US
60/225,268	14 August 2000 (14.08.2000)	US
60/225,758	14 August 2000 (14.08.2000)	US
60/225,267	14 August 2000 (14.08.2000)	US
60/225,214	14 August 2000 (14.08.2000)	US
60/226,279	18 August 2000 (18.08.2000)	US
60/226,868	22 August 2000 (22.08.2000)	US
60/227,182	22 August 2000 (22.08.2000)	US
60/226,681	22 August 2000 (22.08.2000)	US
60/227,009	23 August 2000 (23.08.2000)	US
60/228,924	30 August 2000 (30.08.2000)	US
60/229,344	1 September 2000 (01.09.2000)	US
60/229,343	1 September 2000 (01.09.2000)	US
60/229,287	1 September 2000 (01.09.2000)	US
60/229,345	1 September 2000 (01.09.2000)	US
60/229,513	5 September 2000 (05.09.2000)	US
60/229,509	5 September 2000 (05.09.2000)	US
60/230,438	6 September 2000 (06.09.2000)	US
60/230,437	6 September 2000 (06.09.2000)	US
60/231,413	8 September 2000 (08.09.2000)	US
60/232,080	8 September 2000 (08.09.2000)	US
60/231,414	8 September 2000 (08.09.2000)	US
60/231,244	8 September 2000 (08.09.2000)	US
60/232,081	8 September 2000 (08.09.2000)	US
60/231,242	8 September 2000 (08.09.2000)	US
60/231,243	8 September 2000 (08.09.2000)	US
60/231,968	12 September 2000 (12.09.2000)	US
60/232,401	14 September 2000 (14.09.2000)	US
60/232,399	14 September 2000 (14.09.2000)	US
60/232,400	14 September 2000 (14.09.2000)	US
60/232,397	14 September 2000 (14.09.2000)	US
60/233,063	14 September 2000 (14.09.2000)	US
60/233,064	14 September 2000 (14.09.2000)	US
60/233,065	14 September 2000 (14.09.2000)	US
60/232,398	14 September 2000 (14.09.2000)	US
60/234,223	21 September 2000 (21.09.2000)	US
60/234,274	21 September 2000 (21.09.2000)	US
60/234,997	25 September 2000 (25.09.2000)	US
60/234,998	25 September 2000 (25.09.2000)	US
60/235,484	26 September 2000 (26.09.2000)	US
60/235,834	27 September 2000 (27.09.2000)	US
60/235,836	27 September 2000 (27.09.2000)	US
60/236,369	29 September 2000 (29.09.2000)	US

[Continued on next page]

(54) Title: NUCLEIC ACIDS, PROTEINS, AND ANTIBODIES

(57) Abstract: The present invention relates to novel immune/hematopoietic-related polynucleotides and the polypeptides encoded by these polynucleotides herein collectively known as "immune/hematopoietic antigens", and the use of such immune/hematopoietic antigens for detecting immune/hematopoietic-related diseases and/or disorders, particularly the presence of cancer and cancer metastases of cells of hematopoietic origin. More specifically, isolated immune/hematopoietic associated nucleic acid molecules are provided encoding novel immune/hematopoietic associated polypeptides. Novel immune/hematopoietic polypeptides and antibodies that bind to these polypeptides are provided. Also provided are vectors, host cells, and recombinant and synthetic methods for producing human immune/hematopoietic associated polynucleotides and/or polypeptides. The invention further relates to diagnostic and therapeutic methods useful for diagnosing, treating, preventing and/or prognosing disorders related to the immune system or cells and tissues associated with hematopoiesis, including cancers of cells of hematopoietic origin, and therapeutic methods for treating such disorders. The invention further relates to screening methods for identifying agonists and antagonists of polynucleotides and polypeptides of the invention. The present invention further relates to methods and/or compositions for inhibiting the production and function of the polypeptides of the present invention.

WO 01/57182 A2

60/236,327	29 September 2000 (29.09.2000)	US
60/236,370	29 September 2000 (29.09.2000)	US
60/236,368	29 September 2000 (29.09.2000)	US
60/236,367	29 September 2000 (29.09.2000)	US
60/237,039	2 October 2000 (02.10.2000)	US
60/237,038	2 October 2000 (02.10.2000)	US
60/237,040	2 October 2000 (02.10.2000)	US
60/237,037	2 October 2000 (02.10.2000)	US
60/236,802	2 October 2000 (02.10.2000)	US
60/239,937	13 October 2000 (13.10.2000)	US
60/239,935	13 October 2000 (13.10.2000)	US
60/241,785	20 October 2000 (20.10.2000)	US
60/241,809	20 October 2000 (20.10.2000)	US
60/240,960	20 October 2000 (20.10.2000)	US
60/241,787	20 October 2000 (20.10.2000)	US
60/241,808	20 October 2000 (20.10.2000)	US
60/241,221	20 October 2000 (20.10.2000)	US
60/241,786	20 October 2000 (20.10.2000)	US
60/241,826	20 October 2000 (20.10.2000)	US
60/244,617	1 November 2000 (01.11.2000)	US
60/246,474	8 November 2000 (08.11.2000)	US
60/246,532	8 November 2000 (08.11.2000)	US
60/246,476	8 November 2000 (08.11.2000)	US
60/246,526	8 November 2000 (08.11.2000)	US
60/246,475	8 November 2000 (08.11.2000)	US
60/246,525	8 November 2000 (08.11.2000)	US
60/246,528	8 November 2000 (08.11.2000)	US
60/246,527	8 November 2000 (08.11.2000)	US
60/246,477	8 November 2000 (08.11.2000)	US
60/246,611	8 November 2000 (08.11.2000)	US
60/246,610	8 November 2000 (08.11.2000)	US
60/246,613	8 November 2000 (08.11.2000)	US
60/246,609	8 November 2000 (08.11.2000)	US
60/246,478	8 November 2000 (08.11.2000)	US
60/246,524	8 November 2000 (08.11.2000)	US
60/246,523	8 November 2000 (08.11.2000)	US
60/249,299	17 November 2000 (17.11.2000)	US
60/249,210	17 November 2000 (17.11.2000)	US
60/249,216	17 November 2000 (17.11.2000)	US
60/249,217	17 November 2000 (17.11.2000)	US
60/249,211	17 November 2000 (17.11.2000)	US
60/249,215	17 November 2000 (17.11.2000)	US
60/249,218	17 November 2000 (17.11.2000)	US
60/249,208	17 November 2000 (17.11.2000)	US
60/249,213	17 November 2000 (17.11.2000)	US
60/249,212	17 November 2000 (17.11.2000)	US
60/249,207	17 November 2000 (17.11.2000)	US
60/249,245	17 November 2000 (17.11.2000)	US
60/249,244	17 November 2000 (17.11.2000)	US
60/249,297	17 November 2000 (17.11.2000)	US
60/249,214	17 November 2000 (17.11.2000)	US
60/249,264	17 November 2000 (17.11.2000)	US
60/249,209	17 November 2000 (17.11.2000)	US
60/249,300	17 November 2000 (17.11.2000)	US
60/249,265	17 November 2000 (17.11.2000)	US
60/250,391	1 December 2000 (01.12.2000)	US
60/250,160	1 December 2000 (01.12.2000)	US
60/256,719	5 December 2000 (05.12.2000)	US
60/251,030	5 December 2000 (05.12.2000)	US
60/251,988	5 December 2000 (05.12.2000)	US
60/251,479	6 December 2000 (06.12.2000)	US
60/251,869	8 December 2000 (08.12.2000)	US
60/251,856	8 December 2000 (08.12.2000)	US
60/251,868	8 December 2000 (08.12.2000)	US
60/251,990	8 December 2000 (08.12.2000)	US
60/251,989	8 December 2000 (08.12.2000)	US
60/254,097	11 December 2000 (11.12.2000)	US
60/259,678	5 January 2001 (05.01.2001)	US

(71) **Applicant** *(for all designated States except US)*: **HUMAN GENOME SCIENCES, INC.** [US/US]; 9410 Key West Avenue, Rockville, MD 20850 (US).

(72) **Inventors; and**

(75) **Inventors/Applicants** *(for US only)*: **ROSEN, Craig, A.** [US/US]; 22400 Rolling Hill Lane, Laytonsville, MD 20882 (US). **BARASH, Steven, C.** [US/US]; 111 Watkins Pond Boulevard #301, Rockville, MD 20850 (US). **RUBEN, Steven, M.** [US/US]; 18528 Heritage Hills Drive, Olney, MD 20832 (US).

(74) **Agents: HOOVER, Kenley, K.** et al.; Human Genome Sciences, Inc., 9410 Key West Avenue, Rockville, MD 20850 (US).

(81) **Designated States** *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) **Designated States** *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

Nucleic Acids, Proteins, and Antibodies

[001] This application refers to a "Sequence Listing" that is provided only on electronic media in computer readable form pursuant to Administrative Instructions Section 801(a)(i). The Sequence Listing forms a part of this description pursuant to Rule 5.2 and Administrative Instructions Sections 801 to 806, and is hereby incorporated in its entirety.

[002] The Sequence Listing is provided as an electronic file (PC004PCT_seqList.txt, 76,977,474 bytes in size, created on January 16, 2001) on four identical compact discs (CD-R), labeled "COPY 1," "COPY 2," "COPY 3," and "CRF." The Sequence Listing complies with Annex C of the Administrative Instructions, and may be viewed, for example, on an IBM-PC machine running the MS-Windows operating system by using the V viewer software, version 2000 (see World Wide Web URL: http://www.fileviewer.com).

Field of the Invention

[003] The present invention relates to novel immune system and hematopoietic related (herein "immune/hematopoietic") polynucleotides, the polypeptides encoded by these polynucleotides herein collectively referred to as "immune/hematopoietic antigens," and antibodies that immunospecifically bind these polypeptides, and the use of such immune/hematopoietic polynucleotides, antigens, and antibodies for detecting, treating, preventing and/or prognosing disorders of the immune system, including, but not limited to, the presence of cancer and cancer metastases of cells of hematopitic origin. More specifically, isolated immune/hematopoietic nucleic acid molecules are provided encoding novel immune/hematopoietic polypeptides. Novel

immune/hematopoietic polypeptides and antibodies that bind to these polypeptides are provided. Also provided are vectors, host cells, and recombinant and synthetic methods for producing human immune/hematopoietic polynucleotides, polypeptides, and/or antibodies. The invention further relates to diagnostic and therapeutic methods useful for diagnosing, treating, preventing and/or prognosing disorders related to hematopoiesis and the immune system, including cancers of cells of hematopoietic origins, and therapeutic methods for treating such disorders. The invention further relates to screening methods for identifying agonists and antagonists of polynucleotides and polypeptides of the invention. The invention further relates to methods and/or compositions for inhibiting or promoting the production and/or function of the polypeptides of the invention.

Background of the Invention

[004] The immune system is an intricate network composed of cells, tissues and soluble substances that function to protect the body from invasion by foreign substances and pathogens. The major cells of the immune system are white blood cells, including lymphocytes, such as B cells and T cells, and myeloid cells, such as basophils, eosinophils, neutrophils, mast cells, monocytes, macrophages and dendritic cells. The soluble components of the immune system, are molecules (often polypeptides) that are not contained within cells, but rather are found in extracellular fluids such as lymph and blood plasma. Some of the major soluble substances are antibodies, complement proteins, and cytokines.

[005] Cells of the immune system (as well as red blood cells and platelets) are derived from a common precursor stem cell by a process known as hematopoiesis. During fetal life hematopoiesis occurs in the liver and spleen, but in the adult, hematopoiesis occurs mainly in bone marrow. The stem cells from which all blood cells are derived proliferate and differentiate into the various blood cell lineages, (e.g., lymphocytes (B or T cells), myeloid cells (basophils, eosinophils, neutrophils, mast cells, macrophages), platelets, or red blood cells) in response to signals received from other cells (e.g., stromal cells) in the bone marrow microenvironment and also from cytokines. Many of the cytokines that promote the growth and differentiation of

hematopoietic stem cells are known as "colony stimulating factors". For example, interleukin-3 (IL-3, and also known as multi-colony stimulating factor) and granulocyte macrophage colony stimulationg factor (GM-CSF), which are released by activated macrophages T cells, stimulate the production of macrophages and granulocytes (myelopoiesis). Stem cell factor (SCF, c-kit ligand) is a growth factor for primitive lymphoid and myeloid hematopoietic bone marrow progenitor cells expressing the early cell surface marker CD34. Other hematopoeitic cytokines/growth factors include, but are not limited to macrophage colony stimulating factor (M-CSF) and granulocyte colony stimulating factor (G-CSF). Interleukins-1, 6, and 7 have also been shown to function as hematopoietic growth factors/cytokines.

[006] The maturation of lymphocytes has an added layer of complexity in that each individual T and B cell generates a unique antigen specific receptor – a B cell receptor (antibody) in the case of B cells or a T cell receptor in the case of T cells . Because it is possible that B and T cells may generate autoreactive antigen receptors, B and T cells undergo negative selection processes that eliminate autoreactive lymphocytes from the circulating pool of mature lymphocytes. Defects in negative selection may contribute to the occurrence of autoimmune disease. In addition, T cells undergo a process of positive selection in which T cells are selected for their ability to interact with the major histocompatibility antigens. In the thymus, T cells also differentiate into one of two classes, CD4+ T helper (Th) cells or CD8+ cytotoxic T cells. The majority of the maturation and selection procees occurs in the bone marrow for B cells, whereas T cell progenitor cells migrate from the bone marrow to the thymus where they complete their maturation.

[007] Cells of the immune system circulate throughout the body in both the lymph and the blood. Immune cells will leave the circulatory system and enter the tissues by a process known as diapedesis. Immune cells return to the circulatory system via travel in the lymph. Situated along the lymphatic vessels are lymph nodes, which are small nodular aggregates of lymphoid tissues. The architecture of the lymph node is designed to facilitate acquired immune responses, with antigen presenting cells, B cells and T cells all in close proximity. Antigen presenting cells (APCs, e.g., dendritic cells, macrophages, B cells) display antigen on their surface in the form of peptides associated MHC class II molecules to T helper cells. T helper cells with T-cell

receptors specific for the given antigen become activated if they bind to the peptide MHC complexes and receive co-stimulatory signals (e.g, stimulation of CD28 on the Tcell by B7 molecules on the APC). Activated T helper cells proliferate, secreted cytokines, and can stimulate antigen-specific B cells or T cells to become activated. Once activated, cytotoxic T cells proliferate and are able to induce apoptosis of cells expressing specific antigen on their surface as a peptide in the context of MHC Class I molecules. Activated B cells also proliferate and may either enter into germinal center and undergo a process of affinity maturation of their antigen receptor, or differentiate into antibody forming cells (plasma cells) that secrete large quantities of antigen-specific antibody.

[008] Aside from lymphocytes and antigen presenting cells, introduced above, there are several other accessory cells in the immune system including neutrophils, eosinophils, basophils, mast cells, and Natural Killer (NK) cells. NK cells are large granular lymphocytes that have cytotoxic function, especially against cells infected with intracellular pathogens, and may function in the eradication of cancer cells. Neutrophils are phagocytic cells that play a key role in the inflammatory process. Activated mast cells release granules containing histamine and other active agents which are effective against large parasites and also contribute to allergic reactions and asthma. Eosinophils bear Fc receptors for IgG and IgE, and participate in the killing of antibody coated parasites.

[009] The immune system can be classified into the acquired and innate immune system. The cells of the innate immune system (e.g., neutrophils, eosinophils, basophils, mast cells) are not antigen specific and their action is not enhanced by repeated exposure to the same antigen. The cells of the acquired immune system (B and T cells) are antigen specific and repeated exposure of B and T cells to an antigen results in improved immune repsonses (memory responses) produced by these cell types. The cells and products of the acquired immune system can function to focus the action of the innate immune system. For example, eosinophils are not in themselves antigen specific, but as a result of expression of Fc receptors on their surface, their activity can be focused on a specific antigen to which an antibody response has been made by the acquired immune system. For a more extensive review of the immune

system, see *Fundamental Immunology* , 4th edition, ed. William Paul, Lippincott-Raven Pub. (1998).

[010] As illustrated above, an immune response is seldom carried out by a single cell type, but rather requires the coordinated efforts of several cell types. In order to coordinate an immune response, it is necessary that cells of the immune system communicate with each other and with other cells of the body. Communication between cells may be made by cell-cell contact, between membrane bound molecules on each cell, or by the interaction of soluble components of the immune system with cellular receptors. Usually, such receptors are embedded in the plasma membrane, but there also exist a subset of cytoplasmic and nuclear receptors. Communication, or signaling, between cell types may have one or more of a variety of consequences including, activation, proliferation, differentiation, or apoptosis. Activation and differentiation may result in the expression or secretion of polypeptides, or other molecules, which in turn affect the function of other cells and/or molecules of the immune system.

[011] Signaling molecules of the immune system, including not only cellular receptors and ligands, but also the downstream effectors of the receptors and/or ligands, may be described as immunomodulators. In addition, immunomodulators (also known as biological response modifiers) include microbial or synthetic substances and products of activated cells. The mechanism of action of immunomodulators usually involves a complicated interplay of various regulator and effector systems. Immunomodulators may enhance (immunoprophylaxis, immunostimulation), restore (immunosubstitution, immunorestoration) or suppress (immunosuppression, immunodeviation) immunological functions or activities. Immunomodulators may be, for example, cytokines, cytokine receptors, inhibitors of DNA synthesis, intacellular receptors, or components of signal transduction pathways, some of which are described in more detail below:

Cytokines and Cytokine Receptors

[012] Cytokines are small soluble proteins produced by one cell that alter the behavior or other properties of another cell or itself. Thus, by definition, cytokines are immunomodulatory molecules. Many cytokines have multiple biological effects and

are critical to the regulation of the immune response. For a review on cytokines, refer to Chapter 11 of *Cellular and Molecular Immunology* by Abbas et al. (1991).

[013] Immune responses of the acquired immune system can be classified into two broad classes of immune responses: humoral (antibody-mediated) immune responses and cell-mediated immune responses (cell-mediated, i.e., cytotoxic T cell, immune response). Both types of responses require activation of CD4+ T helper cells. Depending on several factors, of which one factor is the cytokine environment, T helper (Th) cells may differentiate into either Th1 cells that promote cell-mediated responses or Th2 cells that promote humoral responses. Th1 cells, which produce interferon (IFN)-gamma, interleukin (IL)-2 and tumor necrosis factor (TNF)-beta, evoke cell-mediated immunity and phagocyte-dependent inflammation. Th2 cells, which produce IL-4, IL-5, IL-6, IL-9, IL-10, and IL-13, evoke strong antibody responses (including those of the IgE class) and eosinophil accumulation, but inhibit several functions of phagocytic cells (phagocyte-independent inflammation). The presence of Th1 or Th2 T cells can have a dramatic effect on the outcome of infection. A Th1 response during the course of infection by the intracellular bacterium *mycobacterium leprae (M. leprae)* is protective, whereas a Th2 response is much less so. Patients that make Th2 response to M. leprae develop full-blown lepromatous leprosy which is eventually fatal. The (mis)regulation of Th1 and Th2 responses have been implicated in the pathogenesis of several diseases, including several organ-specific autoimmune disorders such as Crohn's disease, sarcoidosis, acute kidney allograft rejection, some unexplained recurrent abortions. For a review on Th1 and Th2 subsets, see Romagnani, *Ann. Allergy Asthma Immunol.* 85:9-18 (2000).

[014] From the preceding example it is apparent that cytokines have play key roles on the class and effectiveness of the immune response. It is important to note that cytokines have effects on cell of both the innate and acquired immune systems and are produced by both immune and non-immune cells types.

[015] Other cytokines such as interferon-alpha (secreted by leukocytes) and interferon–beta (secreted by fibroblasts and many other cell types) are cytokines that function to target the immune system towards fighting viral infections. The binding of interferon-alpha and -beta to cells results in a cellular signalling cascade which ultimately results in the inhibition of viral replication in infected cells, the upregulation

of MHC class I expression on cells, and the activation of Natural Killer (NK) cells. Interferons are useful in the diagnosis, treatment and prevention of viral infections and cancers.

Intracellular immunomodulators.

[016] Immunomodulatory proteins are not only cytokines or cytokine receptors. They may also be located intracellularly. For, example they may be intracellular components of a signaling pathway, or even intracellular receptors for certain signaling molecules such as steroids. One example of intracellular immunomodulatory proteins are the immunophilins such as cyclophilin and FK binding protein (FKBP). These immunophilins are peptidyl-prolyl cis-trans isomerases, though their enzymatic ability may be distinct from their role as immunomodulators. When these molecules are bound by the drugs, Cyclosporin A and FK506, respectively, they in turn inhibit the action of activated calcineurin. Calcineurin is a calcium activated serine/threonine kinase which dephosphorylates the transcription factor Nuclear Factor of Activated T cells (NF-AT). Upon dephosphorylation, NF-AT enters the nucleus and induces the transcription of several genes including IL-2. In sum, the immunophilin:drug complexes are able to inhibit clonal expansion of T cells by inhibiting IL-2 synthesis. In addition, FKBP when bound to another drug, rapamycin, can also inhibit the signaling of IL-2 through the IL-2 receptor. FKBP:rapamycin complexes accomplish the inhibition of IL-2 signaling not by binding to calcineurin, but by binding to and inactivating the protein kinases associated with IL-2 signaling resulting in the same outcome, the inhibition of T cell clonal expansion.

[017] Defects in any one or more of the components of the immune system can lead to diseaseor susceptibility to infectious diseases. Two major classes of immune system disorders are autoimmune diseases, and immunodeficiencies. In autoimmunity, the effector mechanisms of the immune system (e.g., antigen specific antibodies and cellular cytotoxicity, e.g., of cytotoxic T cells, or natural killer cells) are misdirected at self rather than foreign antigens resulting is tissue distruction. Diseases classified as or associated with immunodeficiencies are diseases in which the immune system is unable to mount an effective immune response. A classic example of an immunodeficiecy is X-linked agammaglobulinemia in which an intracellular

signalling molecule expressed in B lymphocytes (Bruton's tyrosine kinase) is defective. The loss of function of this kinase prevents B cell maturation, thus patients with X linked agammaglobulinemia do not have mature B cells and are unable to make antibody, and as a result are susceptible to infection.

[018] The discovery of new human immune/hematopoietic polynucleotides, the polypeptides encoded by them, and antibodies that immunospecifically bind these polypeptides, satisfies a need in the art by providing new compositions which are useful in the diagnosis, treatment, prevention and/or prognosis of disorders of the immune system, including, but not limited to, autoimmune disorders, (e.g., systemic lupus erythematosus, rheumatoid arthritis, idiopathic thrombocytopenic purpura and multiple sclerosis) and immunodeficiencies (e.g., X-linked agammaglobulinemia, severe combined immunodeficiency, Wiskott-Aldrich syndrome, and ataxia telangiectasia). Additionally, immune/hematopoietic molecules would be useful as agents to boost immune responsiveness to pathogens or to suppress immune reactions, for example as is necessary in conjunction with organ transplantation.

Summary of the Invention

[019] The present invention relates to novel immune/hematopoietic related polynucleotides, the polypeptides encoded by these polynucleotides herein collectively referred to as "immune/hematopoietic antigens," and antibodies that immunospecifically bind these polypeptides, and the use of such immune/hematopoietic polynucleotides, antigens, and antibodies for detecting, treating, preventing and/or prognosing disorders of the immune system, including, but not limited to, the presence of cancer and cancer metastases of cells of hematopoietic origin. More specifically, isolated immune/hematopoietic nucleic acid molecules are provided encoding novel immune/hematopoietic polypeptides. Novel immune/hematopoietic polypeptides and antibodies that bind to these polypeptides are provided. Also provided are vectors, host cells, and recombinant and synthetic methods for producing human immune/hematopoietic polynucleotides, polypeptides, and/or antibodies. The invention further relates to diagnostic and therapeutic methods useful for diagnosing, treating, preventing and/or prognosing disorders related to the immune system or hematopoitic cells or tisues, including cancers of cells of

hematopoietic origin, and therapeutic methods for treating such disorders. The invention further relates to screening methods for identifying agonists and antagonists of polynucleotides and polypeptides of the invention. The invention further relates to methods and/or compositions for inhibiting or promoting the production and/or function of the polypeptides of the invention.

Detailed Description

Tables

[020] Table 1A summarizes some of the polynucleotides encompassed by the invention (including cDNA clones related to the sequences (Clone ID NO:Z), contig sequences (contig identifier (Contig ID:) and contig nucleotide sequence identifier (SEQ ID NO:X)) and further summarizes certain characteristics of these polynucleotides and the polypeptides encoded thereby. The first column provides a unique clone identifier, "Clone ID NO:Z", for a cDNA plasmid related to each immune/hematopoietic associated contig sequence disclosed in Table 1A. The second column provides a unique contig identifier, "Contig ID:" for each of the contig sequences disclosed in Table 1A. The third column provides the sequence identifier, "SEQ ID NO:X", for each of the contig polynucleotide sequences disclosed in Table 1A. The fourth column, "ORF (From-To)", provides the location (i.e., nucleotide position numbers) within the polynucleotide sequence of SEQ ID NO:X that delineate the preferred open reading frame (ORF) shown in the sequence listing and referenced in Table 1A as SEQ ID NO:Y (column 5). Column 6 lists residues comprising predicted epitopes contained in the polypeptides encoded by each of the preferred ORFs (SEQ ID NO:Y). Identification of potential immunogenic regions was performed according to the method of Jameson and Wolf (CABIOS, 4:181-186 (1988)); specifically, the Genetics Computer Group (GCG) implementation of this algorithm, embodied in the program PEPTIDESTRUCTURE (Wisconsin Package v10.0, Genetics Computer Group (GCG), Madison, Wisc.). This method returns a measure of the probability that a given residue is found on the surface of the protein. Regions where the antigenic index score is greater than 0.9 over at least 6 amino acids are indicated in Table 1A as "Predicted Epitopes." In particular embodiments, immune/hematopoietic associated polypeptides of the invention comprise, or

alternatively consist of, one, two, three, four, five or more of the predicted epitopes described in Table 1A. It will be appreciated that depending on the analytical criteria used to predict antigenic determinants, the exact address of the determinant may vary slightly. Column 7, "Tissue Distribution" shows the expression profile of tissue, cells, and/or cell line libraries which express the polynucleotides of the invention. The first number in column 7 (preceding the colon), represents the tissue/cell source identifier code corresponding to the code and description provided in Table 4. Expression of these polynucleotides was not observed in the other tissues and/or cell libraries tested. For those identifier codes in which the first two letters are not "AR", the second number in column 7 (following the colon) represents the number of times a sequence corresponding to the reference polynucleotide sequence (e.g., SEQ ID NO:X) was identified in the tissue/cell source. Those tissue/cell source identifier codes in which the first two letters are "AR" designate information generated using DNA array technology. Utilizing this technology, cDNAs were amplified by PCR and then transferred, in duplicate, onto the array. Gene expression was assayed through hybridization of first strand cDNA probes to the DNA array. cDNA probes were generated from total RNA extracted from a variety of different tissues and cell lines. Probe synthesis was performed in the presence of ^{33}P dCTP, using oligo(dT) to prime reverse transcription. After hybridization, high stringency washing conditions were employed to remove non-specific hybrids from the array. The remaining signal, emanating from each gene target, was measured using a Phosphorimager. Gene expression was reported as Phosphor Stimulating Luminescence (PSL) which reflects the level of phosphor signal generated from the probe hybridized to each of the gene targets represented on the array. A local background signal subtraction was performed before the total signal generated from each array was used to normalize gene expression between the different hybridizations. The value presented after "[array code]:" represents the mean of the duplicate values, following background subtraction and probe normalization. One of skill in the art could routinely use this information to identify normal and/or diseased tissue(s) which show a predominant expression pattern of the corresponding polynucleotide of the invention or to identify polynucleotides which show predominant and/or specific tissue and/or cell expression. Column 8, "Cytologic Band," provides the chromosomal location of polynucleotides

corresponding to SEQ ID NO:X. Chromosomal location was determined by finding exact matches to EST and cDNA sequences contained in the NCBI (National Center for Biotechnology Information) UniGene database. Given a presumptive chromosomal location, disease locus association was determined by comparison with the Morbid Map, derived from Online Mendelian Inheritance in Man (Online Mendelian Inheritance in Man, OMIM™. McKusick-Nathans Institute for Genetic Medicine, Johns Hopkins University (Baltimore, MD) and National Center for Biotechnology Information, National Library of Medicine (Bethesda, MD) 2000. World Wide Web URL: http://www.ncbi.nlm.nih.gov/omim/). If the putative chromosomal location of the Query overlapped with the chromosomal location of a Morbid Map entry, an OMIM identification number is provided in Table 1A, column 9 labeled "OMIM Disease Reference(s)". A key to the OMIM reference identification numbers is provided in Table 5.

[021] Table 1B summarizes additional polynucleotides encompassed by the invention (including cDNA clones related to the sequences (Clone ID NO:Z), contig sequences (contig identifier (Contig ID:) contig nucleotide sequence identifiers (SEQ ID NO:X)), and genomic sequences (SEQ ID NO:B). The first column provides a unique clone identifier, "Clone ID NO:Z", for a cDNA clone related to each contig sequence. The second column provides the sequence identifier, "SEQ ID NO:X", for each contig sequence. The third column provides a unique contig identifier, "Contig ID:" for each contig sequence. The fourth column, provides a BAC identifier "BAC ID NO:A" for the BAC clone referenced in the corresponding row of the table. The fifth column provides the nucleotide sequence identifier, "SEQ ID NO:B" for a fragment of the BAC clone identified in column four of the corresponding row of the table. The sixth column, "Exon From-To", provides the location (i.e., nucleotide position numbers) within the polynucleotide sequence of SEQ ID NO:B which delineate certain polynucleotides of the invention that are also exemplary members of polynucleotide sequences that encode polypeptides of the invention (e.g., polypeptides containing amino acid sequences encoded by the polynucleotide sequences delineated in column six, and fragments and variants thereof).

[022] Table 2 summarizes homology and features of some of the polypeptides of the invention. The first column provides a unique clone identifier, "Clone ID NO:Z",

corresponding to a cDNA disclosed in Table 1A. The second column provides the unique contig identifier, "Contig ID:" corresponding to contigs in Table 1A and allowing for correlation with the information in Table 1A. The third column provides the sequence identifier, "SEQ ID NO:X", for the contig polynucleotide sequences. The fourth column provides the analysis method by which the homology/identity disclosed in the row was determined. Comparisons were made between polypeptides encoded by the polynucleotides of the invention and either a non-redundant protein database (herein referred to as "NR"), or a database of protein families (herein referred to as "PFAM") as further described below. The fifth column provides a description of PFAM/NR hits having significant matches to a polypeptide of the invention. Column six provides the accession number of the PFAM/NR hit disclosed in the fifth column. Column seven, "Score/Percent Identity", provides a quality score or the percent identity, of the hit disclosed in column five. Columns 8 and 9, "NT From" and "NT To" respectively, delineate the polynucleotides in "SEQ ID NO:X" that encode a polypeptide having a significant match to the PFAM/NR database as disclosed in the fifth column. In specific embodiments, polypeptides of the invention comprise, or alternatively consist of, an amino acid sequence encoded by the polynucleotides in SEQ ID NO:X as delineated in columns 8 and 9, or fragments or variants thereof.

[023] Table 3 provides polynucleotide sequences that may be disclaimed according to certain embodiments of the invention. The first column provides a unique clone identifier, "Clone ID NO:Z", for a cDNA clone related to immune/hematopoietic associated contig sequences disclosed in Table 1A. The second column provides the sequence identifier, "SEQ ID NO:X", for contig polynucleotide sequences disclosed in Table 1A. The third column provides the unique contig identifier, "Contig ID", for contigs disclosed in Table 1A. The fourth column provides a unique integer 'a' where 'a' is any integer between 1 and the final nucleotide minus 15 of SEQ ID NO:X, represented as "Range of a", and the fifth column provides a unique integer 'b' where 'b' is any integer between 15 and the final nucleotide of SEQ ID NO:X, represented as "Range of b", where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:X, and where b is greater than or equal to a + 14. For each of the polynucleotides shown as SEQ ID NO:X, the uniquely defined integers can be substituted into the general formula of a-b, and used to describe polynucleotides which

may be preferably excluded from the invention. In certain embodiments, preferably excluded from the polynucleotides of the invention (including polynucleotide fragments and variants as described herein and diagnostic and/or therapeutic uses based on these polynucleotides) are at least one, two, three, four, five, ten, or more of the polynucleotide sequence(s) having the accession number(s) disclosed in the sixth column of this Table (including for example, published sequence in connection with a particular BAC clone). In further embodiments, preferably excluded from the invention are the specific polynucleotide sequence(s) contained in the clones corresponding to at least one, two, three, four, five, ten, or more of the available material having the accession numbers identified in the sixth column of this Table (including for example, the actual sequence contained in an identified BAC clone).

[024] Table 4 provides a key to the tissue/cell source identifier code disclosed in Table 1A, column 7. Column 1 provides the key to the tissue/cell source identifier code disclosed in Table 1A, Column 7. Columns 2-5 provide a description of the tissue or cell source. Codes corresponding to diseased tissues are indicated in column 6 with the word "disease". The use of the word "disease" in column 6 is non-limiting. The tissue or cell source may be specific (e.g. a neoplasm), or may be disease-associated (e.g., a tissue sample from a normal portion of a diseased organ). Furthermore, tissues and/or cells lacking the "disease" designation may still be derived from sources directly or indirectly involved in a disease state or disorder, and therefore may have a further utility in that disease state or disorder. In numerous cases where the tissue/cell source is a library, column 7 identifies the vector used to generate the library.

[025] Table 5 provides a key to the OMIM™ reference identification numbers disclosed in Table 1A, column 9. OMIM reference identification numbers (Column 1) were derived from Online Mendelian Inheritance in Man (Online Mendelian Inheritance in Man, OMIM™. McKusick-Nathans Institute for Genetic Medicine, Johns Hopkins University (Baltimore, MD) and National Center for Biotechnology Information, National Library of Medicine, (Bethesda, MD) 2000. World Wide Web URL: http://www.ncbi.nlm.nih.gov/omim/). Column 2 provides diseases associated with the cytologic band disclosed in Table 1A, column 8, as determined from the Morbid Map database.

[026] Table 6 summarizes ATCC Deposits, Deposit dates, and ATCC designation numbers of deposits made with the ATCC in connection with the present application.

[027] Table 7 shows the cDNA libraries sequenced, tissue source description, vector information and ATCC designation numbers relating to these cDNA libraries.

[028] Table 8 provides a physical characterization of clones encompassed by the invention. The first column provides the unique clone identifier, "Clone ID NO:Z", for certain cDNA clones of the invention, as described in Table 1A. The second column provides the size of the cDNA insert contained in the corresponding cDNA clone.

Definitions

[029] The following definitions are provided to facilitate understanding of certain terms used throughout this specification.

[030] In the present invention, "isolated" refers to material removed from its original environment (e.g., the natural environment if it is naturally occurring), and thus is altered "by the hand of man" from its natural state. For example, an isolated polynucleotide could be part of a vector or a composition of matter, or could be contained within a cell, and still be "isolated" because that vector, composition of matter, or particular cell is not the original environment of the polynucleotide. The term "isolated" does not refer to genomic or cDNA libraries, whole cell total or mRNA preparations, genomic DNA preparations (including those separated by electrophoresis and transferred onto blots), sheared whole cell genomic DNA preparations or other compositions where the art demonstrates no distinguishing features of the polynucleotide sequences of the present invention.

[031] As used herein, a "polynucleotide" refers to a molecule having a nucleic acid sequence encoding SEQ ID NO:Y or a fragment or variant thereof, a nucleic acid sequence contained in SEQ ID NO:X (as described in column 3 of Table 1A) or the complement thereof, a cDNA sequence contained in Clone ID NO:Z (as described in column 1 of Table 1A and contained within a library deposited with the ATCC); a nucleotide sequence encoding the polypeptide encoded by a nucleotide sequence in SEQ ID NO:B as defined in column 6 of Table 1B or a fragment or variant thereof; or a nucleotide coding sequence in SEQ ID NO:B as defined in column 6 of Table 1B or

the complement thereof. For example, the polynucleotide can contain the nucleotide sequence of the full length cDNA sequence, including the 5' and 3' untranslated sequences, the coding region, as well as fragments, epitopes, domains, and variants of the nucleic acid sequence. Moreover, as used herein, a "polypeptide" refers to a molecule having an amino acid sequence encoded by a polynucleotide of the invention as broadly defined (obviously excluding poly-Phenylalanine or poly-Lysine peptide sequences which result from translation of a polyA tail of a sequence corresponding to a cDNA).

[032] As used herein, a "immune/hematopoietic antigen" refers collectively to any polynucleotide disclosed herein (e.g., a nucleic acid sequence contained in SEQ ID NO:X or the complement therof, or cDNA sequence contained in Clone ID NO:Z, or a nucleotide sequence encoding the polypeptide encoded by a nucleotide sequence in SEQ ID NO:B as defined in column 6 of Table 1B, or a nucleotide coding sequence in SEQ ID NO:B as defined in column 6 of Table 1B or the complement thereof and fragments or variants thereof as described herein) or any polypeptide disclosed herein (e.g., an amino acid sequence contained in SEQ ID NO:Y, an amino acid sequence encoded by SEQ ID NO:X, or the complement thereof, an amino acid sequence encoded by the cDNA sequence contained in Clone ID NO:Z, an amino acid sequence encoded by SEQ ID NO:B, or the complement thereof, and fragments or variants thereof as described herein). These immune/hematopoietic antigens have been determined to be predominantly expressed in hematopoietic tissues (e.g., bone marrow, fetal liver, and fetal spleen) or cells and tissues of the immune system (e.g., lymph nodes, spleen, B cells, T cells, monocytes, macrophages, dendritic cells, neutrophils, mast cells, basophils, and eosinophils) including normal or diseased tissues (as shown in Table 1A column 7 and Table 4).

[033] In the present invention, "SEQ ID NO:X" was often generated by overlapping sequences contained in multiple clones (contig analysis). A representative clone containing all or most of the sequence for SEQ ID NO:X is deposited at Human Genome Sciences, Inc. (HGS) in a catalogued and archived library. As shown, for example, in column 1 of Table 1A, each clone is identified by a cDNA Clone ID (identifier generally referred to herein as Clone ID NO:Z). Each Clone ID is unique to an individual clone and the Clone ID is all the information needed to retrieve a given

clone from the HGS library. Furthermore, certain clones disclosed in this application have been deposited with the ATCC on October 5, 2000, having the ATCC designation numbers PTA 2574 and PTA 2575; and on January 5, 2001, having the depositor reference numbers TS-1, TS-2, AC-1, and AC-2. In addition to the individual cDNA clone deposits, most of the cDNA libraries from which the clones were derived were deposited at the American Type Culture Collection (hereinafter "ATCC"). Table 7 provides a list of the deposited cDNA libraries. One can use the Clone ID NO:Z to determine the library source by reference to Tables 6 and 7. Table 7 lists the deposited cDNA libraries by name and links each library to an ATCC Deposit. Library names contain four characters, for example, "HTWE." The name of a cDNA clone (Clone ID NO:Z) isolated from that library begins with the same four characters, for example "HTWEP07". As mentioned below, Table 1A correlates the Clone ID NO:Z names with SEQ ID NO:X. Thus, starting with an SEQ ID NO:X, one can use Tables 1A, 6 and 7 to determine the corresponding Clone ID NO:Z, which library it came from and which ATCC deposit the library is contained in. Furthermore, it is possible to retrieve a given cDNA clone from the source library by techniques known in the art and described elsewhere herein. The ATCC is located at 10801 University Boulevard, Manassas, Virginia 20110-2209, USA. The ATCC deposits were made pursuant to the terms of the Budapest Treaty on the international recognition of the deposit of microorganisms for the purposes of patent procedure.

[034] In specific embodiments, the polynucleotides of the invention are at least 15, at least 30, at least 50, at least 100, at least 125, at least 500, or at least 1000 continuous nucleotides but are less than or equal to 300 kb, 200 kb, 100 kb, 50 kb, 15 kb, 10 kb, 7.5 kb, 5 kb, 2.5 kb, 2.0 kb, or 1 kb, in length. In a further embodiment, polynucleotides of the invention comprise a portion of the coding sequences, as disclosed herein, but do not comprise all or a portion of any intron. In another embodiment, the polynucleotides comprising coding sequences do not contain coding sequences of a genomic flanking gene (i.e., 5' or 3' to the gene of interest in the genome). In other embodiments, the polynucleotides of the invention do not contain the coding sequence of more than 1000, 500, 250, 100, 50, 25, 20, 15, 10, 5, 4, 3, 2, or 1 genomic flanking gene(s).

[035] A "polynucleotide" of the present invention also includes those polynucleotides capable of hybridizing, under stringent hybridization conditions, to sequences contained in SEQ ID NO:X, or the complement thereof (e.g., the complement of any one, two, three, four, or more of the polynucleotide fragments described herein), the polynucleotide sequence delineated in columns 8 and 9 of Table 2 or the complement thereof, and/or cDNA sequences contained in Clone ID NO:Z (e.g., the complement of any one, two, three, four, or more of the polynucleotide fragments, or the cDNA clone within the pool of cDNA clones deposited with the ATCC, described herein) and/or the polynucleotide sequence delineated in column 6 of Table 1B or the complement thereof. "Stringent hybridization conditions" refers to an overnight incubation at 42 degree C in a solution comprising 50% formamide, 5x SSC (750 mM NaCl, 75 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 µg/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1x SSC at about 65 degree C.

[036] Also contemplated are nucleic acid molecules that hybridize to the polynucleotides of the present invention at lower stringency hybridization conditions. Changes in the stringency of hybridization and signal detection are primarily accomplished through the manipulation of formamide concentration (lower percentages of formamide result in lowered stringency), salt conditions, or temperature. For example, lower stringency conditions include an overnight incubation at 37 degree C in a solution comprising 6X SSPE (20X SSPE = 3M NaCl; 0.2M NaH_2PO_4; 0.02M EDTA, pH 7.4), 0.5% SDS, 30% formamide, 100 ug/ml salmon sperm blocking DNA; followed by washes at 50 degree C with 1XSSPE, 0.1% SDS. In addition, to achieve even lower stringency, washes performed following stringent hybridization can be done at higher salt concentrations (e.g. 5X SSC).

[037] Note that variations in the above conditions may be accomplished through the inclusion and/or substitution of alternate blocking reagents used to suppress background in hybridization experiments. Typical blocking reagents include Denhardt's reagent, BLOTTO, heparin, denatured salmon sperm DNA, and commercially available proprietary formulations. The inclusion of specific blocking reagents may require modification of the hybridization conditions described above, due to problems with compatibility.

[038] Of course, a polynucleotide which hybridizes only to polyA+ sequences (such as any 3' terminal polyA+ tract of a cDNA shown in the sequence listing), or to a complementary stretch of T (or U) residues, would not be included in the definition of "polynucleotide," since such a polynucleotide would hybridize to any nucleic acid molecule containing a poly (A) stretch or the complement thereof (e.g., practically any double-stranded cDNA clone generated using oligo dT as a primer).

[039] The polynucleotide of the present invention can be composed of any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. For example, polynucleotides can be composed of single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, the polynucleotide can be composed of triple-stranded regions comprising RNA or DNA or both RNA and DNA. A polynucleotide may also contain one or more modified bases or DNA or RNA backbones modified for stability or for other reasons. "Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications can be made to DNA and RNA; thus, "polynucleotide" embraces chemically, enzymatically, or metabolically modified forms.

[040] The polypeptide of the present invention can be composed of amino acids joined to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres, and may contain amino acids other than the 20 gene-encoded amino acids. The polypeptides may be modified by either natural processes, such as posttranslational processing, or by chemical modification techniques which are well known in the art. Such modifications are well described in basic texts and in more detailed monographs, as well as in a voluminous research literature. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Polypeptides may be branched, for example, as a result of

ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched, and branched cyclic polypeptides may result from posttranslation natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphotidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cysteine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, pegylation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination. (See, for instance, PROTEINS - STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993); POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic Press, New York, pgs. 1-12 (1983); Seifter et al., Meth. Enzymol. 182:626-646 (1990); Rattan et al., Ann. N.Y. Acad. Sci. 663:48-62 (1992).)

[041] "SEQ ID NO:X" refers to a polynucleotide sequence described, for example, in Tables 1A or 2, while "SEQ ID NO:Y" refers to a polypeptide sequence described in column 5 of Table 1A. SEQ ID NO:X is identified by an integer specified in column 3 of Table 1A. The polypeptide sequence SEQ ID NO:Y is a translated open reading frame (ORF) encoded by polynucleotide SEQ ID NO:X. "Clone ID NO:Z" refers to a cDNA clone described in column 1 of Table 1A.

[042] "A polypeptide having biological activity" refers to a polypeptide exhibiting activity similar to, but not necessarily identical to, an activity of a polypeptide of the present invention, including mature forms, as measured in a particular biological assay, with or without dose dependency. In the case where dose dependency does exist, it need not be identical to that of the polypeptide, but rather substantially similar to the dose-dependence in a given activity as compared to the polypeptide of the present invention (i.e., the candidate polypeptide will exhibit greater activity or not more than about 25-fold less and, preferably, not more than about tenfold less activity,

and most preferably, not more than about three-fold less activity relative to the polypeptide of the present invention).

[043] Table 1A summarizes some of the immune/hematopoietic associated polynucleotides encompassed by the invention (including contig sequences (SEQ ID NO:X) and clones (Clone ID NO:Z) and further summarizes certain characteristics of these polynucleotides and the polypeptides encoded thereby.

Polynucleotides and Polypeptides

TABLE 1A

Clone ID NO: Z	Contig ID:	SEQ ID NO: X	ORF (From-To)	AA SEQ ID NO: Y	Predicted Epitopes	Tissue Distribution Library code: count (see Table IV for Library Codes)	Cytologic Band	OMIM Disease Reference(s):
HAMHB21	961376	11	17 - 2389	9763		AR089: 3, AR061: 1 H0271: 18, H0556: 9, H0265: 8, H0581: 8, L0761: 4, H0543: 4, H0422: 4, H0656: 3, H0457: 3, L0766: 3, T0002: 2, L0748: 2, H0220: 1, H0650: 1, S0282: 1, H0610: 1, H0069: 1, H0635: 1, H0179: 1, H0416: 1, H0031: 1, H0090: 1, T0041: 1, H0560: 1, H0529: 1, L0667: 1, L0649: 1, L0803: 1, L0659: 1, L0666: 1, S0052: 1, S0216: 1, H0702: 1, H0518: 1, H0521: 1, L0750: 1, H0445: 1, H0423: 1, H0677: 1 and H0506: 1.		
HASAX16	573692	12	138 - 302	9764	Tyr-20 to Ile-33.	H0004: 2		
HASAY74	526312	13	79 - 318	9765		H0004: 1, H0090: 1 and H0543: 1.		
HASAY89	958768	14	125 - 280	9766	Glu-28 to Lys-34.	H0004: 1, L0749: 1 and H0445: 1.		
HASAY94	521835	15	2 - 103	9767	Thr-10 to Asn-22.	H0004: 2		
HBCAL36	931477	16	3 - 440	9768	Glu-26 to Ser-31, Glu-36 to Gln-55, Glu-70 to Asn-82, Ser-93 to Pro-98, Ser-115 to Ser-121.	H0370: 1, H0581: 1 and H0264: 1.		
HBCAL39	964871	17	333 - 491	9769	Arg-1 to Ser-6.	H0255: 2, H0370: 2 and S0053: 1.		
HBCAM74	864366	18	1 - 348	9770	Pro-29 to Ser-37.	H0341: 1, H0370: 1 and		

HYAAJ03	923526	6872	148 - 279	16624	Glu-26 to Cys-31.	H0583: 1 and L0777: 1.		
HYAAF78	773433	6873	1 - 219	16625	Arg-17 to Asn-23, Ser-43 to Ile-48.	H0583: 1, L0748: 1 and L0754: 1.		
HYAAD24	677019	6874	2 - 157	16626		H0583: 1 and L0748: 1.		
HYAAC24	677010	6875	186 - 389	16627	Glu-21 to Tyr-31, Lys-58 to Lys-64.	H0583: 1 and L0362: 1.		
HYAAB40	841969	6876	268 - 510	16628	Ile-15 to Ser-20.	H0583: 1 and L0717: 1.		
HWWEW11	965434	6877	298 - 471	16629	Gly-1 to Gly-29.	H0657: 1 and L0664: 1.		
HWWEL11	965426	6878	204 - 374	16630	Ser-1 to Leu-8.	H0657: 1 and L0367: 1.		
HWWEC06	933690	6879	85 - 234	16631	Ser-33 to Gly-43.	L0766: 2, H0657: 1 and L0749: 1.		
HWBGG82	779081	6880	377 - 538	16632		H0580: 1, L0745: 1 and L0600: 1.		
HWBGF04	927284	6881	1 - 489	16633		H0580: 1 and L0750: 1.		
HWBFY46	800638	6882	1 - 270	16634	Arg-5 to Arg-14.	AR054: 2, AR050: 2, AR051: 1 H0580: 1		
HWBEG18	909798	6883	55 - 696	16635	Trp-46 to Lys-51, Pro-109 to Asn-123, Phe-156 to Gly-165.	AR089: 82, AR061: 18 H0580: 1		
HWBEF69	754328	6884	324 - 229	16636		H0580: 1 and L0748: 1.		
HWBEE25	677660	6885	534 - 316	16637		H0580: 1 and L0742: 1.		
HWBEA06	934820	6886	31 - 237	16638	Gln-26 to Ser-31, Pro-34 to Gln-40.	H0580: 1 and L0766: 1.		
HWBDU71	835888	6887	418 - 531	16639		H0580: 1 and L0523: 1.		
HWBDU58	735774	6888	490 - 257	16640	Ser-18 to Asn-23, Thr-30 to Thr-41, Pro-63 to Arg-75.	L0439: 3 and H0580: 1.		
HWBDS59	743462	6889	274 - 447	16641	Pro-11 to Thr-20.	L0749: 2, H0580: 1 and L0747: 1.	22q11.2-qter	
HWBDS12	971872	6890	3 - 200	16642	Ile-17 to Glu-37.	H0580: 1, L0775: 1 and L0755: 1.		
HWBDR82	779002	6891	153 - 437	16643	Phe-1 to Arg-13, Ser-21 to Gln-32.	H0580: 1 and L0748: 1.		
HWBDQ58	735769	6892	553 - 744	16644		L0748: 2 and H0580: 1.		
HWBDL82	728805	6893	26 - 355	16645	Pro-64 to Ile-75,	L0439: 2 and H0580: 1.		

[044] The first column in Table 1A provides a unique "Clone ID NO:Z" for a cDNA clone related to each contig sequence disclosed in Table 1A. This clone ID references the cDNA clone which contains at least the 5' most sequence of the assembled contig and at least a portion of SEQ ID NO:X was determined by directly sequencing the referenced clone. The reference clone may have more sequence than described in the sequence listing or the clone may have less. In the vast majority of cases, however, the clone is believed to encode a full-length polypeptide. In the case where a clone is not full-length, a full-length cDNA can be obtained by methods known in the art and/or as described elsewhere herein.

[045] The second column in Table 1A provides a unique "Contig ID" identification for each contig sequence. The third column provides the "SEQ ID NO:X" identifier for each of the immune/hematopoietic associated contig polynucleotide sequences disclosed in Table 1A. The fourth column, "ORF (From-To)", provides the location (i.e., nucleotide position numbers) within the polynucleotide sequence "SEQ ID NO:X" that delineate the preferred open reading frame (ORF) shown in the sequence listing and referenced in Table 1A, column 5, as SEQ ID NO:Y. Where the nucleotide position number "To" is lower than the nucleotide position number "From", the preferred ORF is the reverse complement of the referenced polynucleotide sequence.

[046] The fifth column in Table 1A provides the corresponding SEQ ID NO:Y for the polypeptide sequence encoded by the preferred ORF delineated in column 4. In one embodiment, the invention provides an amino acid sequence comprising, or alternatively consisting of, a polypeptide encoded by the portion of SEQ ID NO:X delineated by "ORF (From-To)". Also provided are polynucleotides encoding such amino acid sequences and the complementary strand thereto.

[047] Column 6 in Table 1A lists residues comprising epitopes contained in the polypeptides encoded by the preferred ORF (SEQ ID NO:Y), as predicted using the algorithm of Jameson and Wolf, (1988) Comp. Appl. Biosci. 4:181-186. The Jameson-Wolf antigenic analysis was performed using the computer program PROTEAN (Version 3.11 for the Power MacIntosh, DNASTAR, Inc., 1228 South Park Street Madison, WI). In preferred embodiments, polypeptides of the invention comprise, or alternatively consist of, at least one, two, three, four, five or more of the predicted epitopes as described in Table 1A. It will be appreciated that depending on the

TABLE 1B

Clone ID NO:Z	SEQ ID NO:X	CONTIG ID:	BAC ID: A	SEQ ID NO:B	EXON From-To
HAMHB21	11	961376	AL035530	19515	1-137 465-2035 2761-2939 3499-3663 3841-4296 5031-5501 5707-5905 5946-6330 6790-6899 6957-7401 7628-7818 7889-10548
HAMHB21	11	961376	AL035530	19516	1-577
HASAX16	12	573692	AL358796	19517	1-1065
HASAX16	12	573692	AL356513	19518	1-1281
HASAX16	12	573692	AL358796	19519	1-964
HASAX16	12	573692	AL356513	19520	1-540
HASAX16	12	573692	AL356513	19521	1-967
HASAY74	13	526312	AC006479	19522	1-1081
HASAY74	13	526312	AC006479	19523	1-1052
HASAY74	13	526312	AC006479	19524	1-117
HASAY89	14	958768	AC012580	19525	1-874
HASAY89	14	958768	AL133502	19526	1-874
HASAY89	14	958768	AC012580	19527	1-87
HASAY89	14	958768	AC012580	19528	1-1034 1426-2177
HASAY89	14	958768	AL133502	19529	1-1034 1426-2177
HASAY94	15	521835	AC022702	19530	1-220 2365-2620 2992-3310 3432-3987 4545-5235 5266-5325 5814-6468 6801-6965 7327-7763 7979-8172
HASAY94	15	521835	AC025796	19531	1-272
HASAY94	15	521835	AC024998	19532	1-272
HASAY94	15	521835	AL365438	19533	1-272
HASAY94	15	521835	AL390122	19534	1-272
HASAY94	15	521835	AL133216	19535	1-306 1488-1890 2450-2722 3082-3400 3526-4081 4304-5314
HASAY94	15	521835	AF198096	19536	1-308 1490-1896 2457-2728 2816-3407 3532-4087

HWWEL11	6878	965426	AL352978	35897	1-139
HWWEL11	6878	965426	AL359402	35898	1-686
HWWEL11	6878	965426	AL359402	35899	1-139
HWWEC06	6879	933690	AC013443	35900	1-1665
HWWEC06	6879	933690	AC013443	35901	1-574
HWBGG82	6880	779081	AC027208	35902	1-680
HWBGG82	6880	779081	AC027208	35903	1-209
HWBGG82	6880	779081	AC027208	35904	1-284
HWBFY46	6882	800638	AL157838	35905	1-509 799-2464 3520-3635 4689-6422 7888-7978 9634-10252 10983-11450 12705-14906
HWBFY46	6882	800638	AL157838	35906	1-341
HWBFY46	6882	800638	AL157838	35907	1-566
HWBEF69	6884	754328	AF127936	35908	1-864
HWBEF69	6884	754328	AF127936	35909	1-763
HWBEF69	6884	754328	AF127936	35910	1-857
HWBEE25	6885	677660	AL358975	35911	1-123 577-1935 2158-2318 2782-3217 3421-3612
HWBEE25	6885	677660	AL358975	35912	1-590
HWBEE25	6885	677660	AL358975	35913	1-514
HWBEA06	6886	934820	AL162427	35914	1-1679
HWBEA06	6886	934820	AL162427	35915	1-1109
HWBEA06	6886	934820	AL162427	35916	1-370 997-1347 3930-4323 5372-5564
HWBDS59	6889	743462	Z71183	35917	1-1948 2434-2823 3625-3940 5198-7203 7399-7681 7762-7863 8123-8772
HWBDS59	6889	743462	Z71183	35918	1-370
HWBDS59	6889	743462	Z71183	35919	1-547
HWBDR82	6891	779002	AC024998	35920	1-345 926-1432
HWBDR82	6891	779002	AC016992	35921	1-168 636-891 1001-1260 1329-1668 2257-2762 2790-3281 3615-4100 4815-5257 5264-5728
HWBDR82	6891	779002	AF198096	35922	1-168 636-891

[052] Table 1B summarizes additional polynucleotides encompassed by the invention (including cDNA clones related to the sequences (Clone ID NO:Z), contig sequences (contig identifier (Contig ID:) contig nucleotide sequence identifiers (SEQ ID NO:X)), and genomic sequences (SEQ ID NO:B). The first column provides a unique clone identifier, "Clone ID NO:Z", for a cDNA clone related to each contig sequence. The second column provides the sequence identifier, "SEQ ID NO:X", for each contig sequence. The third column provides a unique contig identifier, "Contig ID:" for each contig sequence. The fourth column, provides a BAC identifier "BAC ID NO:A" for the BAC clone referenced in the corresponding row of the table. The fifth column provides the nucleotide sequence identifier, "SEQ ID NO:B" for a fragment of the BAC clone identified in column four of the corresponding row of the table. The sixth column, "Exon From-To", provides the location (i.e., nucleotide position numbers) within the polynucleotide sequence of SEQ ID NO:B which delineate certain polynucleotides of the invention that are also exemplary members of polynucleotide sequences that encode polypeptides of the invention (e.g., polypeptides containing amino acid sequences encoded by the polynucleotide sequences delineated in column six, and fragments and variants thereof).

TABLE 2

Clone ID NO:Z	Contig ID:	SEQ ID NO:X	Analysis Method	PFam/NR Description	PFam/NR Accession Number	Score/ Percent Identity	NT From	NT To
HAMHB21	961376	11	HMMER 2.1.1	PFAM: RhoGAP domain	PF00620	100	497	916
HBDAC79	935414	26	HMMER 2.1.1	PFAM: Acyl-CoA dehydrogenase	PF00441	102.6	94	342
			blastx.14	(AL021958) fadE9 [Mycobacterium tuberculosis]	gi|2911026|emb|CAA 17519.1|	62% 51% 64% 35%	94 250 348 6	255 384 422 89
HBJAG72	722723	32	HMMER 1.8	PFAM: Zinc-binding metalloprotease domain	PF00099	2.32	83	115
HBJGT92	919507	107	HMMER 1.8	PFAM: Bacterial murT protein	PF00293	7.68	377	436
HBJIY20	669519	141	HMMER 2.1.1	PFAM: Orn/Lys/Arg decarboxylase	PF01276	53.1	49	162
HBMBU24	677240	234	HMMER 1.8	PFAM: Core histones H2A, H2B, H3 and H4	PF00125	7.72	139	192
HBMBZ71	880580	238	HMMER 1.8	PFAM: von Willebrand factor type D domain	PF00094	26.52	200	526
HBMDC16	888206	258	blastx.2	Butyrophilin-like protein BUTR-1.	sp|AAF72554|AAF72 554	50%	19	315
HBMUO90	928078	294	blastx.14	Zfp61p [Mus musculus]	gi|887887|gb|AAC52 290.1|	84% 53%	24 154	119 243
HBMXE31	573323	317	HMMER 1.8	PFAM: Gonadotropin-releasing hormones	PF00446	10.78	112	141
HCFCF47	894415	372	HMMER 1.8	PFAM: Eukaryotic protein kinase domain	PF00069	89.54	20	295
HCFCJ21	671028	374	HMMER	PFAM: Core histones	PF00125	12.79	123	188

						37%	282	211
HYABP53	971448	6848	HMMER 2.1.1	PFAM: Immunoglobulin domain	PF00047	27.8	257	427
HYABB43	800006	6860	HMMER 2.1.1	PFAM: Alpha-2-macroglobulin family	PF00207	90.9	206	27
			blastx.2	alpha-2-macroglobulin [Mus musculus]	gb|AAA39508.1|	67%	263	36
HYAAU65	909956	6863	HMMER 2.1.1	PFAM: RasGEF domain	PF00617	47	2	112
HWBEG18	909798	6883	HMMER 2.1.1	PFAM: EF hand	PF00036	33.3	505	591
			blastx.14	(AF106071) ras activator RasGRP [Homo sapiens]	gi|4038292|gb|AAC97349.1|	71% 59% 46% 70% 48% 66%	698 382 103 604 2 619	889 597 420 684 142 654
HWBCS08	909031	6897	HMMER 2.1.1	PFAM: B-box zinc finger.	PF00643	32.5	263	388
			blastx.14	(AF119043) transcriptional intermediary factor 1 gamma; TIF1gamma [Homo sapiens]	gi|4325109|gb|AAD17259.1|	92%	2	562
HWBAQ71	944080	6918	blastx.2	predicted using Genefinder [Caenorhabditis elegans]	emb|CAA97423.1|	34%	141	464
HWAFG12	969605	6934	blastx.14	(AF071081) proline-rich mucin homolog [Mycobacterium tuberculosis]	gi|5305335|gb|AAD41594.1|AF071081_1	34% 50% 28% 40%	112 195 398 365	17 142 264 300

[053] Table 2 further characterizes certain encoded polypeptides of the invention, by providing the results of comparisons to protein and protein family databases. The first column provides a unique clone identifier, "Clone ID NO:", corresponding to a cDNA clone disclosed in Table 1A. The second column provides the unique contig indentifier, "Contig ID:" which allows correlation with the information in Table 1A. The third column provides the sequence identifier, "SEQ ID NO:X", for the contig polynucleotide sequences. The fourth column provides the analysis method by which the homology/identity disclosed in the row was determined. The fifth column provides a description of PFam/NR hits having significant matches identified by each analysis. Column six provides the accession number of the PFam/NR hit disclosed in the fifth column. Column seven, "Score/Percent Identity", provides a quality score or the percent identity, of the hit disclosed in column five. Comparisons were made between polypeptides encoded by polynucleotides of the invention and a non-redundant protein database (herein referred to as "NR"), or a database of protein families (herein referred to as "PFam"), as described below.

[054] The NR database, which comprises the NBRF PIR database, the NCBI GenPept database, and the SIB SwissProt and TrEMBL databases, was made non-redundant using the computer program nrdb2 (Warren Gish, Washington University in Saint Louis). Each of the polynucleotides shown in Table 1A, column 3 (e.g., SEQ ID NO:X or the 'Query' sequence) was used to search against the NR database. The computer program BLASTX was used to compare a 6-frame translation of the Query sequence to the NR database (for information about the BLASTX algorithm please see Altshul et al., J. Mol. Biol. 215:403-410 (1990), and Gish et al., Nat. Genet. 3:266-272 (1993)). A description of the sequence that is most similar to the Query sequence (the highest scoring 'Subject') is shown in column five of Table 2 and the database accession number for that sequence is provided in column six. The highest scoring 'Subject' is reported in Table 2 if (a) the estimated probability that the match occurred by chance alone is less than 1.0e-07, and (b) the match was not to a known repetitive element. BLASTX returns alignments of short polypeptide segments of the Query and Subject sequences which share a high degree of similarity; these segments are known as High-Scoring Segment Pairs or HSPs. Table 2 reports the degree of similarity between the Query and the Subject for each HSP as a percent identity in Column 7. The percent

identity is determined by dividing the number of exact matches between the two aligned sequences in the HSP, dividing by the number of Query amino acids in the HSP and multiplying by 100. The polynucleotides of SEQ ID NO:X which encode the polypeptide sequence that generates an HSP are delineated by columns 8 and 9 of Table 2.

[055] The PFam database, PFam version 5.2, (Sonnhammer et al., Nucl. Acids Res., 26:320-322, (1998)) consists of a series of multiple sequence alignments; one alignment for each protein family. Each multiple sequence alignment is converted into a probability model called a Hidden Markov Model, or HMM, that represents the position-specific variation among the sequences that make up the multiple sequence alignment (see, e.g., R. Durbin et al., *Biological sequence analysis: probabilistic models of proteins and nucleic acids*, Cambridge University Press, 1998 for the theory of HMMs). The program HMMER version 1.8 (Sean Eddy, Washington University in Saint Louis) was used to compare the predicted protein sequence for each Query sequence (SEQ ID NO:Y in Table 1A) to each of the HMMs derived from PFam version 5.2. A HMM derived from PFam version 5.2 was said to be a significant match to a polypeptide of the invention if the score returned by HMMER 1.8 was greater than 0.8 times the HMMER 1.8 score obtained with the most distantly related known member of that protein family. The description of the PFam family which shares a significant match with a polypeptide of the invention is listed in column 5 of Table 2, and the database accession number of the PFam hit is provided in column 6. Column 7 provides the score returned by HMMER version 1.8 for the alignment. Columns 8 and 9 delineate the polynucleotides of SEQ ID NO:X which encode the polypeptide sequence which shows a significant match to a PFam protein family.

[056] As mentioned, columns 8 and 9 in Table 2, "NT From" and "NT To", delineate the polynucleotides of "SEQ ID NO:X" that encode a polypeptide having a significant match to the PFam/NR database as disclosed in the fifth column of Table 2. In one embodiment, the invention provides a protein comprising, or alternatively consisting of, a polypeptide encoded by the polynucleotides of SEQ ID NO:X delineated in columns 8 and 9 of Table 2. Also provided are polynucleotides encoding such proteins, and the complementary strand thereto.

[057] The nucleotide sequence SEQ ID NO:X and the translated SEQ ID NO:Y are sufficiently accurate and otherwise suitable for a variety of uses well known in the art and described further below. For instance, the nucleotide sequences of SEQ ID NO:X are useful for designing nucleic acid hybridization probes that will detect nucleic acid sequences contained in SEQ ID NO:X or the cDNA contained in Clone ID NO:Z. These probes will also hybridize to nucleic acid molecules in biological samples, thereby enabling immediate applications in chromosome mapping, linkage analysis, tissue identification and/or typing, and a variety of forensic and diagnostic methods of the invention. Similarly, polypeptides identified from SEQ ID NO:Y may be used to generate antibodies which bind specifically to these polypeptides, or fragments thereof, and/or to the polypeptides encoded by the cDNA clones identified in, for example, Table 1A.

[058] Nevertheless, DNA sequences generated by sequencing reactions can contain sequencing errors. The errors exist as misidentified nucleotides, or as insertions or deletions of nucleotides in the generated DNA sequence. The erroneously inserted or deleted nucleotides cause frame shifts in the reading frames of the predicted amino acid sequence. In these cases, the predicted amino acid sequence diverges from the actual amino acid sequence, even though the generated DNA sequence may be greater than 99.9% identical to the actual DNA sequence (for example, one base insertion or deletion in an open reading frame of over 1000 bases).

[059] Accordingly, for those applications requiring precision in the nucleotide sequence or the amino acid sequence, the present invention provides not only the generated nucleotide sequence identified as SEQ ID NO:X, and a predicted translated amino acid sequence identified as SEQ ID NO:Y, but also a sample of plasmid DNA containing cDNA Clone ID NO:Z (deposited with the ATCC on October 5, 2000, and receiving ATCC designation numbers PTA 2574 and PTA 2575; deposited with the ATCC on January 5, 2001, having the depositor reference numbers TS-1, TS-2, AC-1, and AC-2; and/or as set forth, for example, in Table 1A, 6 and 7). The nucleotide sequence of each deposited clone can readily be determined by sequencing the deposited clone in accordance with known methods. Further, techniques known in the art can be used to verify the nucleotide sequences of SEQ ID NO:X.

[060] The predicted amino acid sequence can then be verified from such deposits. Moreover, the amino acid sequence of the protein encoded by a particular clone can also be directly determined by peptide sequencing or by expressing the protein in a suitable host cell containing the deposited human cDNA, collecting the protein, and determining its sequence.

RACE Protocol For Recovery of Full-Length Genes

[061] Partial cDNA clones can be made full-length by utilizing the rapid amplification of cDNA ends (RACE) procedure described in Frohman, M.A., et al., Proc. Nat'l. Acad. Sci. USA, 85:8998-9002 (1988). A cDNA clone missing either the 5' or 3' end can be reconstructed to include the absent base pairs extending to the translational start or stop codon, respectively. In some cases, cDNAs are missing the start codon of translation. The following briefly describes a modification of this original 5' RACE procedure. Poly A+ or total RNA is reverse transcribed with Superscript II (Gibco/BRL) and an antisense or complementary primer specific to the cDNA sequence. The primer is removed from the reaction with a Microcon Concentrator (Amicon). The first-strand cDNA is then tailed with dATP and terminal deoxynucleotide transferase (Gibco/BRL). Thus, an anchor sequence is produced which is needed for PCR amplification. The second strand is synthesized from the dA-tail in PCR buffer, Taq DNA polymerase (Perkin-Elmer Cetus), an oligo-dT primer containing three adjacent restriction sites (XhoI, SalI and ClaI) at the 5' end and a primer containing just these restriction sites. This double-stranded cDNA is PCR amplified for 40 cycles with the same primers as well as a nested cDNA-specific antisense primer. The PCR products are size-separated on an ethidium bromide-agarose gel and the region of gel containing cDNA products the predicted size of missing protein-coding DNA is removed. cDNA is purified from the agarose with the Magic PCR Prep kit (Promega), restriction digested with XhoI or SalI, and ligated to a plasmid such as pBluescript SKII (Stratagene) at XhoI and EcoRV sites. This DNA is transformed into bacteria and the plasmid clones sequenced to identify the correct protein-coding inserts. Correct 5' ends are confirmed by comparing this sequence with the putatively identified homologue and overlap with the partial cDNA clone. Similar methods known in the art and/or commercial kits are used to amplify and recover 3' ends.

[062] Several quality-controlled kits are commercially available for purchase. Similar reagents and methods to those above are supplied in kit form from Gibco/BRL for both 5' and 3' RACE for recovery of full length genes. A second kit is available from Clontech which is a modification of a related technique, SLIC (single-stranded ligation to single-stranded cDNA), developed by Dumas et al., Nucleic Acids Res., 19:5227-32 (1991). The major differences in procedure are that the RNA is alkaline hydrolyzed after reverse transcription and RNA ligase is used to join a restriction site-containing anchor primer to the first-strand cDNA. This obviates the necessity for the dA-tailing reaction which results in a polyT stretch that is difficult to sequence past.

[063] An alternative to generating 5' or 3' cDNA from RNA is to use cDNA library double-stranded DNA. An asymmetric PCR-amplified antisense cDNA strand is synthesized with an antisense cDNA-specific primer and a plasmid-anchored primer. These primers are removed and a symmetric PCR reaction is performed with a nested cDNA-specific antisense primer and the plasmid-anchored primer.

RNA Ligase Protocol For Generating The 5' or 3' End Sequences To Obtain Full Length Genes

[064] Once a gene of interest is identified, several methods are available for the identification of the 5' or 3' portions of the gene which may not be present in the original cDNA plasmid. These methods include, but are not limited to, filter probing, clone enrichment using specific probes and protocols similar and identical to 5' and 3' RACE. While the full length gene may be present in the library and can be identified by probing, a useful method for generating the 5' or 3' end is to use the existing sequence information from the original cDNA to generate the missing information. A method similar to 5' RACE is available for generating the missing 5' end of a desired full-length gene. (This method was published by Fromont-Racine et al., Nucleic Acids Res., 21(7):1683-1684 (1993)). Briefly, a specific RNA oligonucleotide is ligated to the 5' ends of a population of RNA presumably containing full-length gene RNA transcript. A primer set containing a primer specific to the ligated RNA oligonucleotide and a primer specific to a known sequence of the gene of interest, is used to PCR amplify the 5' portion of the desired full length gene which may then be sequenced and used to generate the full length gene. This method starts with total RNA isolated from the desired source, poly A RNA may be

used but is not a prerequisite for this procedure. The RNA preparation may then be treated with phosphatase if necessary to eliminate 5' phosphate groups on degraded or damaged RNA which may interfere with the later RNA ligase step. The phosphatase, if used, is then inactivated and the RNA is treated with tobacco acid pyrophosphatase in order to remove the cap structure present at the 5' ends of messenger RNAs. This reaction leaves a 5' phosphate group at the 5' end of the cap cleaved RNA which can then be ligated to an RNA oligonucleotide using T4 RNA ligase. This modified RNA preparation can then be used as a template for first strand cDNA synthesis using a gene specific oligonucleotide. The first strand synthesis reaction can then be used as a template for PCR amplification of the desired 5' end using a primer specific to the ligated RNA oligonucleotide and a primer specific to the known sequence of the immune/hematopoietic antigen of interest. The resultant product is then sequenced and analyzed to confirm that the 5' end sequence belongs to the relevant immune/hematopoietic antigen.

[065] The present invention also relates to vectors or plasmids which include such DNA sequences, as well as the use of the DNA sequences. The material deposited with the ATCC (deposited with the ATCC on October 5, 2000, and receiving ATCC designation numbers PTA 2574 and PTA 2575; deposited with the ATCC on January 5, 2001, having the depositor reference numbers TS-1, TS-2, AC-1, and AC-2; and/or as set forth, for example, in Table 1A, 6 and 7) is a mixture of cDNA clones derived from a variety of human tissue and cloned in either a plasmid vector or a phage vector, as shown, for example, in Table 7. These deposits are referred to as "the deposits" herein. The tissues from which some of the clones were derived are listed in Table 7, and the vector in which the corresponding cDNA is contained is also indicated in Table 7. The deposited material includes cDNA clones corresponding to SEQ ID NO:X described, for example, in Table 1A (Clone ID NO:Z). A clone which is isolatable from the ATCC Deposits by use of a sequence listed as SEQ ID NO:X, may include the entire coding region of a human gene or in other cases such clone may include a substantial portion of the coding region of a human gene. Furthermore, although the sequence listing may in some instances list only a portion of the DNA sequence in a clone included in the ATCC Deposits, it is well within the ability of one skilled in the art to sequence the DNA included in a clone contained in the ATCC Deposits by use of a sequence (or portion

thereof) described in, for example Tables 1A or 2 by procedures hereinafter further described, and others apparent to those skilled in the art.

[066] Also provided in Table 7 is the name of the vector which contains the cDNA clone. Each vector is routinely used in the art. The following additional information is provided for convenience.

[067] Vectors Lambda Zap (U.S. Patent Nos. 5,128,256 and 5,286,636), Uni-Zap XR (U.S. Patent Nos. 5,128,256 and 5,286,636), Zap Express (U.S. Patent Nos. 5,128,256 and 5,286,636), pBluescript (pBS) (Short, J. M. et al., *Nucleic Acids Res. 16:*7583-7600 (1988); Alting-Mees, M. A. and Short, J. M., *Nucleic Acids Res. 17:*9494 (1989)) and pBK (Alting-Mees, M. A. et al., *Strategies 5:*58-61 (1992)) are commercially available from Stratagene Cloning Systems, Inc., 11011 N. Torrey Pines Road, La Jolla, CA, 92037. pBS contains an ampicillin resistance gene and pBK contains a neomycin resistance gene. Phagemid pBS may be excised from the Lambda Zap and Uni-Zap XR vectors, and phagemid pBK may be excised from the Zap Express vector. Both phagemids may be transformed into *E. coli* strain XL-1 Blue, also available from Stratagene.

[068] Vectors pSport1, pCMVSport 1.0, pCMVSport 2.0 and pCMVSport 3.0, were obtained from Life Technologies, Inc., P. O. Box 6009, Gaithersburg, MD 20897. All Sport vectors contain an ampicillin resistance gene and may be transformed into *E. coli* strain DH10B, also available from Life Technologies. See, for instance, Gruber, C. E., et al., *Focus 15:*59- (1993). Vector lafmid BA (Bento Soares, Columbia University, New York, NY) contains an ampicillin resistance gene and can be transformed into *E. coli* strain XL-1 Blue. Vector pCR®2.1, which is available from Invitrogen, 1600 Faraday Avenue, Carlsbad, CA 92008, contains an ampicillin resistance gene and may be transformed into *E. coli* strain DH10B, available from Life Technologies. See, for instance, Clark, J. M., Nuc. Acids Res. 16:9677-9686 (1988) and Mead, D. et al., Bio/Technology 9: (1991).

[069] The present invention also relates to the genes corresponding to SEQ ID NO:X, SEQ ID NO:Y, and/or the deposited clone (Clone ID NO:Z). The corresponding gene can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include preparing probes or primers from the disclosed

sequence and identifying or amplifying the corresponding gene from appropriate sources of genomic material.

[070] Also provided in the present invention are allelic variants, orthologs, and/or species homologs. Procedures known in the art can be used to obtain full-length genes, allelic variants, splice variants, full-length coding portions, orthologs, and/or species homologs of immune/hematopoietic associated genes corresponding to SEQ ID NO:X or the complement thereof, polypeptides encoded by SEQ ID NO:X or the complement thereof, and/or the cDNA contained in Clone ID NO:Z, using information from the sequences disclosed herein or the clones deposited with the ATCC. For example, allelic variants and/or species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source for allelic variants and/or the desired homologue.

[071] The polypeptides of the invention can be prepared in any suitable manner. Such polypeptides include isolated naturally occurring polypeptides, recombinantly produced polypeptides, synthetically produced polypeptides, or polypeptides produced by a combination of these methods. Means for preparing such polypeptides are well understood in the art.

[072] The polypeptides may be in the form of the secreted protein, including the mature form, or may be a part of a larger protein, such as a fusion protein (see below). It is often advantageous to include an additional amino acid sequence which contains secretory or leader sequences, pro-sequences, sequences which aid in purification, such as multiple histidine residues, or an additional sequence for stability during recombinant production.

[073] The polypeptides of the present invention are preferably provided in an isolated form, and preferably are substantially purified. A recombinantly produced version of a polypeptide, including the secreted polypeptide, can be substantially purified using techniques described herein or otherwise known in the art, such as, for example, by the one-step method described in Smith and Johnson, Gene 67:31-40 (1988). Polypeptides of the invention also can be purified from natural, synthetic or recombinant sources using techniques described herein or otherwise known in the art, such as, for example, antibodies of the invention raised against the immune/hematopoietic polypeptides of the present invention in methods which are well known in the art.

[074] The present invention provides a polynucleotide comprising, or alternatively consisting of, the nucleic acid sequence of SEQ ID NO:X, and/or the cDNA sequence contained in Clone ID NO:Z. The present invention also provides a polypeptide comprising, or alternatively, consisting of, the polypeptide sequence of SEQ ID NO:Y, a polypeptide encoded by SEQ ID NO:X or a complement thereof, a polypeptide encoded by the cDNA contained in Clone ID NO:Z, and/or the polypeptide sequence encoded by a nucleotide sequence in SEQ ID NO:B as defined in column 6 of Table 1B. Polynucleotides encoding a polypeptide comprising, or alternatively consisting of the polypeptide sequence of SEQ ID NO:Y, a polypeptide encoded by SEQ ID NO:X, a polypeptide encoded by the cDNA contained in Clone ID NO:Z and/or a polypeptide sequence encoded by a nucleotide sequence in SEQ ID NO:B as defined in column 6 of Table 1B are also encompassed by the invention. The present invention further encompasses a polynucleotide comprising, or alternatively consisting of, the complement of the nucleic acid sequence of SEQ ID NO:X, a nucleic acid sequence encoding a polypeptide encoded by the complement of the nucleic acid sequence of SEQ ID NO:X, and/or the cDNA contained in Clone ID NO:Z.

[075] Moreover, representative examples of polynucleotides of the invention comprise, or alternatively consist of, one, two, three, four, five, six, seven, eight, nine, ten, or more of the sequences delineated in Table 1B column 6, or any combination thereof. Additional, representative examples of polynucleotides of the invention comprise, or alternatively consist of, one, two, three, four, five, six, seven, eight, nine, ten, or more of the complementary strand(s) of the sequences delineated in Table 1B column 6, or any combination thereof. In further embodiments, the above-described polynucleotides of the invention comprise, or alternatively consist of, sequences delineated in Table 1B, column 6, and have a nucleic acid sequence which is different from that of the BAC fragment having the sequence disclosed in SEQ ID NO:B (see Table 1B, column 5). In additional embodiments, the above-described polynucleotides of the invention comprise, or alternatively consist of, sequences delineated in Table 1B, column 6, and have a nucleic acid sequence which is different from that published for the BAC clone identified as BAC ID NO:A (see Table 1B, column 4). In additional embodiments, the above-described polynucleotides of the invention comprise, or alternatively consist of, sequences delineated in Table 1B, column 6, and have a nucleic

acid sequence which is different from that contained in the BAC clone identified as BAC ID NO:A (see Table 1B, column 4). Polypeptides encoded by these polynucleotides, other polynucleotides that encode these polypeptides, and antibodies that bind these polypeptides are also encompassed by the invention. Additionally, fragments and variants of the above-described polynucleotides and polypeptides are also encompassed by the invention.

[076] Further, representative examples of polynucleotides of the invention comprise, or alternatively consist of, one, two, three, four, five, six, seven, eight, nine, ten, or more of the sequences delineated in column 6 of Table 1B which correspond to the same Clone ID NO:Z (see Table 1B, column 1), or any combination thereof. Additional, representative examples of polynucleotides of the invention comprise, or alternatively consist of, one, two, three, four, five, six, seven, eight, nine, ten, or more of the complementary strand(s) of the sequences delineated in column 6 of Table 1B which correspond to the same Clone ID NO:Z (see Table 1B, column 1), or any combination thereof. In further embodiments, the above-described polynucleotides of the invention comprise, or alternatively consist of, sequences delineated in column 6 of Table 1B which correspond to the same Clone ID NO:Z (see Table 1B, column 1) and have a nucleic acid sequence which is different from that of the BAC fragment having the sequence disclosed in SEQ ID NO:B (see Table 1B, column 5). In additional embodiments, the above-described polynucleotides of the invention comprise, or alternatively consist of, sequences delineated in column 6 of Table 1B which correspond to the same Clone ID NO:Z (see Table 1B, column 1) and have a nucleic acid sequence which is different from that published for the BAC clone identified as BAC ID NO:A (see Table 1B, column 4). In additional embodiments, the above-described polynucleotides of the invention comprise, or alternatively consist of, sequences delineated in column 6 of Table 1B which correspond to the same Clone ID NO:Z (see Table 1B, column 1) and have a nucleic acid sequence which is different from that contained in the BAC clone identified as BAC ID NO:A (see Table 1B, column 4). Polypeptides encoded by these polynucleotides, other polynucleotides that encode these polypeptides, and antibodies that bind these polypeptides are also encompassed by the invention. Additionally, fragments and variants of the above-described polynucleotides and polypeptides are also encompassed by the invention.

[077] Further, representative examples of polynucleotides of the invention comprise, or alternatively consist of, one, two, three, four, five, six, seven, eight, nine, ten, or more of the sequences delineated in column 6 of Table 1B which correspond to the same contig sequence identifer SEQ ID NO:X (see Table 1B, column 2), or any combination thereof. Additional, representative examples of polynucleotides of the invention comprise, or alternatively consist of, one, two, three, four, five, six, seven, eight, nine, ten, or more of the complementary strand(s) of the sequences delineated in column 6 of Table 1B which correspond to the same contig sequence identifer SEQ ID NO:X (see Table 1B, column 2), or any combination thereof. In further embodiments, the above-described polynucleotides of the invention comprise, or alternatively consist of, sequences delineated in column 6 of Table 1B which correspond to the same contig sequence identifer SEQ ID NO:X (see Table 1B, column 2) and have a nucleic acid sequence which is different from that of the BAC fragment having the sequence disclosed in SEQ ID NO:B (see Table 1B, column 5). In additional embodiments, the above-described polynucleotides of the invention comprise, or alternatively consist of, sequences delineated in column 6 of Table 1B which correspond to the same contig sequence identifer SEQ ID NO:X (see Table 1B, column 2) and have a nucleic acid sequence which is different from that published for the BAC clone identified as BAC ID NO:A (see Table 1B, column 4). In additional embodiments, the above-described polynucleotides of the invention comprise, or alternatively consist of, sequences delineated in column 6 of Table 1B which correspond to the same contig sequence identifer SEQ ID NO:X (see Table 1B, column 2) and have a nucleic acid sequence which is different from that contained in the BAC clone identified as BAC ID NO:A (See Table 1B, column 4). Polypeptides encoded by these polynucleotides, other polynucleotides that encode these polypeptides, and antibodies that bind these polypeptides are also encompassed by the invention. Additionally, fragments and variants of the above-described polynucleotides and polypeptides are also encompassed by the invention.

[078] Moreover, representative examples of polynucleotides of the invention comprise, or alternatively consist of, one, two, three, four, five, six, seven, eight, nine, ten, or more of the sequences delineated in the same row of Table 1B column 6, or any combination thereof. Additional, representative examples of polynucleotides of the

invention comprise, or alternatively consist of, one, two, three, four, five, six, seven, eight, nine, ten, or more of the complementary strand(s) of the sequences delineated in the same row of Table 1B column 6, or any combination thereof. In preferred embodiments, the polynucleotides of the invention comprise, or alternatively consist of, one, two, three, four, five, six, seven, eight, nine, ten, or more of the complementary strand(s) of the sequences delineated in the same row of Table 1B column 6, wherein sequentially delineated sequences in the table (i.e. corresponding to those exons located closest to each other) are directly contiguous in a 5' to 3' orientation. In further embodiments, above-described polynucleotides of the invention comprise, or alternatively consist of, sequences delineated in the same row of Table 1B, column 6, and have a nucleic acid sequence which is different from that of the BAC fragment having the sequence disclosed in SEQ ID NO:B (see Table 1B, column 5). In additional embodiments, the above-described polynucleotides of the invention comprise, or alternatively consist of, sequences delineated in the same row of Table 1B, column 6, and have a nucleic acid sequence which is different from that published for the BAC clone identified as BAC ID NO:A (see Table 1B, column 4). In additional embodiments, the above-described polynucleotides of the invention comprise, or alternatively consist of, sequences delineated in the same row of Table 1B, column 6, and have a nucleic acid sequence which is different from that contained in the BAC clone identified as BAC ID NO:A (see Table 1B, column 4). Polypeptides encoded by these polynucleotides, other polynucleotides that encode these polypeptides, and antibodies that bind these polypeptides are also encompassed by the invention.

[079] In additional specific embodiments, polynucleotides of the invention comprise, or alternatively consist of, one, two, three, four, five, six, seven, eight, nine, ten, or more of the sequences delineated in column 6 of Table 1B, and the polynucleotide sequence of SEQ ID NO:X (e.g., as defined in Table 1B, column 2) or fragments or variants thereof. Polypeptides encoded by these polynucleotides, other polynucleotides that encode these polypeptides, and antibodies that bind these polypeptides are also encompassed by the invention.

[080] In additional specific embodiments, polynucleotides of the invention comprise, or alternatively consist of, one, two, three, four, five, six, seven, eight, nine, ten, or more of the sequences delineated in column 6 of Table 1B which correspond to

the same Clone ID NO:Z (see Table 1B, column 1), and the polynucleotide sequence of SEQ ID NO:X (e.g., as defined in Table 1A or 1B) or fragments or variants thereof. In preferred embodiments, the delineated sequence(s) and polynucleotide sequence of SEQ ID NO:X correspond to the same Clone ID NO:Z. Polypeptides encoded by these polynucleotides, other polynucleotides that encode these polypeptides, and antibodies that bind these polypeptides are also encompassed by the invention.

[081] In further specific embodiments, polynucleotides of the invention comprise, or alternatively consist of, one, two, three, four, five, six, seven, eight, nine, ten, or more of the sequences delineated in the same row of column 6 of Table 1B, and the polynucleotide sequence of SEQ ID NO:X (e.g., as defined in Table 1A or 1B) or fragments or variants thereof. In preferred embodiments, the delineated sequence(s) and polynucleotide sequence of SEQ ID NO:X correspond to the same row of column 6 of Table 1B. Polypeptides encoded by these polynucleotides, other polynucleotides that encode these polypeptides, and antibodies that bind these polypeptides are also encompassed by the invention.

[082] In additional specific embodiments, polynucleotides of the invention comprise, or alternatively consist of a polynucleotide sequence in which the 3' 10 polynucleotides of one of the sequences delineated in column 6 of Table 1B and the 5' 10 polynucleotides of the sequence of SEQ ID NO:X are directly contiguous. Nucleic acids which hybridize to the complement of these 20 contiguous polynucleotides under stringent hybridization conditions or alternatively, under lower stringency conditions, are also encompassed by the invention. Polypeptides encoded by these polynucleotides and/or nucleic acids, other polynucleotides and/or nucleic acids that encode these polypeptides, and antibodies that bind these polypeptides are also encompassed by the invention. Additionally, fragments and variants of the above-described polynucleotides, nucleic acids, and polypeptides are also encompassed by the invention.

[083] In additional specific embodiments, polynucleotides of the invention comprise, or alternatively consist of, a polynucleotide sequence in which the 3' 10 polynucleotides of one of the sequences delineated in column 6 of Table 1B and the 5' 10 polynucleotides of a fragment or variant of the sequence of SEQ ID NO:X are directly contiguous Nucleic acids which hybridize to the complement of these 20 contiguous polynucleotides under stringent hybridization conditions or alternatively, under lower

stringency conditions, are also encompassed by the invention. Polypeptides encoded by these polynucleotides and/or nucleic acids, other polynucleotides and/or nucleic acids encoding these polypeptides, and antibodies that bind these polypeptides are also encompassed by the invention. Additionally, fragments and variants of the above-described polynucleotides, nucleic acids, and polypeptides are also encompassed by the invention.

[084] In specific embodiments, polynucleotides of the invention comprise, or alternatively consist of, a polynucleotide sequence in which the 3' 10 polynucleotides of the sequence of SEQ ID NO:X and the 5' 10 polynucleotides of the sequence of one of the sequences delineated in column 6 of Table 1B are directly contiguous. Nucleic acids which hybridize to the complement of these 20 contiguous polynucleotides under stringent hybridization conditions or alternatively, under lower stringency conditions, are also encompassed by the invention. Polypeptides encoded by these polynucleotides and/or nucleic acids, other polynucleotides and/or nucleic acids encoding these polypeptides, and antibodies that bind these polypeptides are also encompassed by the invention. Additionally, fragments and variants of the above-described polynucleotides, nucleic acids, and polypeptides are also encompassed by the invention.

[085] In specific embodiments, polynucleotides of the invention comprise, or alternatively consist of, a polynucleotide sequence in which the 3' 10 polynucleotides of a fragment or variant of the sequence of SEQ ID NO:X and the 5' 10 polynucleotides of the sequence of one of the sequences delineated in column 6 of Table 1B are directly contiguous. Nucleic acids which hybridize to the complement of these 20 contiguous polynucleotides under stringent hybridization conditions or alternatively, under lower stringency conditions, are also encompassed by the invention. Polypeptides encoded by these polynucleotides and/or nucleic acids, other polynucleotides and/or nucleic acids encoding these polypeptides, and antibodies that bind these polypeptides are also encompassed by the invention. Additionally, fragments and variants of the above-described polynucleotides, nucleic acids, and polypeptides, are also encompassed by the invention.

[086] In further specific embodiments, polynucleotides of the invention comprise, or alternatively consist of, a polynucleotide sequence in which the 3' 10 polynucleotides of one of the sequences delineated in column 6 of Table 1B and the 5' 10

polynucleotides of another sequence in column 6 are directly contiguous. Nucleic acids which hybridize to the complement of these 20 contiguous polynucleotides under stringent hybridization conditions or alternatively, under lower stringency conditions, are also encompassed by the invention. Polypeptides encoded by these polynucleotides and/or nucleic acids, other polynucleotides and/or nucleic acids encoding these polypeptides, and antibodies that bind these polypeptides are also encompassed by the invention. Additionally, fragments and variants of the above-described polynucleotides, nucleic acids, and polypeptides are also encompassed by the invention.

[087] In specific embodiments, polynucleotides of the invention comprise, or alternatively consist of, a polynucleotide sequence in which the 3' 10 polynucleotides of one of the sequences delineated in column 6 of Table 1B and the 5' 10 polynucleotides of another sequence in column 6 corresponding to the same Clone ID NO:Z (see Table 1B, column 1) are directly contiguous. Nucleic acids which hybridize to the complement of these 20 lower stringency conditions, are also encompassed by the invention. Polypeptides encoded by these polynucleotides and/or nucleic acids, other polynucleotides and/or nucleic acids encoding these polypeptides, and antibodies that bind these polypeptides are also encompassed by the invention. Additionally, fragments and variants of the above-described polynucleotides, nucleic acids, and polypeptides are also encompassed by the invention.

[088] In specific embodiments, polynucleotides of the invention comprise, or alternatively consist of, a polynucleotide sequence in which the 3' 10 polynucleotides of one sequence in column 6 corresponding to the same contig sequence identifer SEQ ID NO:X (see Table 1B, column 2) are directly contiguous. Nucleic acids which hybridize to the complement of these 20 contiguous polynucleotides under stringent hybridization conditions or alternatively, under lower stringency conditions, are also encompassed by the invention. Polypeptides encoded by these polynucleotides and/or nucleic acids, other polynucleotides and/or nucleic acids encoding these polypeptides, and antibodies that bind these polypeptides are also encompassed by the invention. Additionally, fragments and variants of the above-described polynucleotides, nucleic acids, and polypeptides are also encompassed by the invention.

[089] In specific embodiments, polynucleotides of the invention comprise, or alternatively consist of a polynucleotide sequence in which the 3' 10 polynucleotides of

one of the sequences delineated in column 6 of Table 1B and the 5' 10 polynucleotides of another sequence in column 6 corresponding to the same row are directly contiguous. In preferred embodiments, the 3' 10 polynucleotides of one of the sequences delineated in column 6 of Table 1B is directly contiguous with the 5' 10 polynucleotides of the next sequential exon delineated in Table 1B, column 6. Nucleic acids which hybridize to the complement of these 20 contiguous polynucleotides under stringent hybridization conditions or alternatively, under lower stringency conditions, are also encompassed by the invention. Polypeptides encoded by these polynucleotides and/or nucleic acids, other polynucleotides and/or nucleic acids encoding these polypeptides, and antibodies that bind these polypeptides are also encompassed by the invention. Additionally, fragments and variants of the above-described polynucleotides, nucleic acids, and polypeptides are also encompassed by the invention.

[090] Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. Accordingly, for each contig sequence (SEQ ID NO:X) listed in the third column of Table 1A, preferably excluded are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 and the final nucleotide minus 15 of SEQ ID NO:X, b is an integer of 15 to the final nucleotide of SEQ ID NO:X, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:X, and where b is greater than or equal to a + 14. More specifically, preferably excluded are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a and b are integers as defined in columns 4 and 5, respectively, of Table 3. In specific embodiments, the polynucleotides of the invention do not consist of at least one, two, three, four, five, ten, or more of the specific polynucleotide sequences referenced by the Genbank Accession No. as disclosed in column 6 of Table 3 (including for example, published sequence in connection with a particular BAC clone). In further embodiments, preferably excluded from the invention are the specific polynucleotide sequence(s) contained in the clones corresponding to at least one, two, three, four, five, ten, or more of the available material having the accession numbers identified in the sixth column of this Table (including for

example, the actual sequence contained in an identified BAC clone). In no way is this listing meant to encompass all of the sequences which may be excluded by the general formula, it is just a representative example. All references available through these accessions are hereby incorporated by reference in their entirety.

TABLE 3

Clone ID NO: Z	SEQ ID NO: X	Contig ID:	EST Disclaimer Range of a	EST Disclaimer Range of b	Accession #'s
HAMHB21	11	961376	1 - 3744	15 - 3758	AI760827, AW408019, AI253155, AI349366, AI356482, AA814034, AW075920, AW407984, AI760691, AA251937, AI766650, AA352825, AA243541, AI934100, AA352840, H72208, H72106, AW193021, AL035530, AL035530, and AL035530.
HASAX16	12	573692	1 - 305	15 - 319	AL356513, AL356513, AL356513, AL358796, and AL358796.
HASAY74	13	526312	1 - 375	15 - 389	AC006479, AC006479, AC006479, and AC006479.
HASAY89	14	958768	1 - 849	15 - 863	AI240823, AC012580, AC012580, AC012580, AL133502, and AL133502.
HASAY94	15	521835	1 - 89	15 - 103	AL133216, AC024998, AC024998, AL365438, AL365438, AL390122, AL390122, AL133216, AL133216, AL133216, AC022702, AC022702, AC025796, AC025796, AF198096, AF198096, and AF198096.
HBCAL36	16	931477	1 - 553	15 - 567	AI908374.
HBCAL39	17	964871	1 - 587	15 - 601	AL117337, AF205588, AL022345, AL022345, AL022345, AL161931, AL161931, AL161931, AL354975, AL354975, AL354975, AL355583, AL355583, AL355583, AL023808, AL023808, AF205588, AF205588, AF205588, AL117337, AL117337, and AL117337.
HBCAM74	18	864366	1 - 437	15 - 451	AA642815.
HBCAR79	19	573989	1 - 415	15 - 429	AC021215.
HBCAS69	20	669802	1 - 274	15 - 288	
HBCAT17	21	503573	1 - 385	15 - 399	AI150061, AA382210, AC011448, and AC011448.
HBCAT63	22	573993	1 - 224	15 - 238	AI114447, AL133243, AL133243, AL133243, and AL133243.
HBCBM52	23	932514	1 - 657	15 - 671	AA312172.
HBCBX12	24	861018	1 - 219	15 - 233	AB023431, AC005954, AC005954, AC005954, and AC068475.
HBCBZ05	25	922800	1 - 713	15 - 727	
HBDAC79	26	935414	1 - 612	15 - 626	AA904376, H23157, AA663798, AA359944, AA448036, R61429, H09004, AA309011, AF126245, AP001785, and AP001785.
HBDAD04	27	614849	1 - 375	15 - 389	AI057188, AW273323, C15735, AI932257, AI028587, AI940493, AW043720, AW021811, AA858079, AW197210, AI791278, AA508647, C15853, AA181800, AW043729, AW008982, AI918922, AI985340, AA031428, AA282285, AC007392, Z82198, AL031121, AL121840, L24095, AC004882, AC005529, AC002070, AC008984, AL132985, AL049781, Z72006, AC005881, AL079305, AP000263, AC005888, AC003080, AC004935, AL035667, AC005610, AP000036, AC005186, AC004842, AL035696, AC003661, AC004503, AC009263, AL023581, AC003009, U41096, AL021368, AP000567, AC004933, AC003975, AP000494, AL008638, AF015262, AL049833, AL034350, AL031010, AJ229043, AP001069, AC006288, AL049561, AL024507, AC005701, AP000100, AC005792, AP000566, AL023283, AC011594, AC006377,

					AF196969, AL109798, M89651, AL031589, and AC004477.
HYACC03	6859	923525	1 - 709	15 - 723	AA179370, and AA464124.
HYABB43	6860	800006	1 - 449	15 - 463	AA131782, AI279457, AI459300, AI887969, and AC010175.
HYAAY57	6861	734463	1 - 430	15 - 444	T66305, AA886512, F09927, AA749328, AA815166, AI475828, AA648819, AA651977, AA953511, and AA741442.
HYAAV45	6862	717479	1 - 530	15 - 544	W88974, and AC004582.
HYAAU65	6863	909956	1 - 581	15 - 595	T81523, T98761, AL080117, AB023176, and U14103.
HYAAT55	6864	731420	1 - 398	15 - 412	T98379.
HYAAT39	6865	929758	1 - 447	15 - 461	R62243, AI298594, H00843, AA400786, AW453087, W69160, AA485564, W57903, AI810717, AI860626, R69650, and R69664.
HYAAN84	6866	782838	1 - 802	15 - 816	AA464831, AA180467, AA505048, AA449184, and AC012083.
HYAAL49	6867	973089	1 - 479	15 - 493	
HYAAL41	6868	887395	1 - 622	15 - 636	AW392670, U46347, AB026436, AR066494, AR060234, A81671, AR054110, and AR069079.
HYAAL19	6869	668338	1 - 922	15 - 936	AA082707, AA081874, AI201257, AA489677, H22947, AI269725, N63194, H11596, F04937, R39853, N63182, R05455, and AC005070.
HYAAK77	6870	887044	1 - 765	15 - 779	R84433, H83979, and AA017024.
HYAAK74	6871	765252	1 - 510	15 - 524	R93968, W93000, and AL109613.
HYAAJ03	6872	923526	1 - 434	15 - 448	AI092745.
HYAAF78	6873	773433	1 - 467	15 - 481	H01888, R93377, AL110203, Z83843, and AC004386.
HYAAD24	6874	677019	1 - 550	15 - 564	H63893.
HYAAC24	6875	677010	1 - 675	15 - 689	AA075368, AA075369, AA559324, AA804735, AI336026, AI114694, AI732327, AC007406, Z69364, D87009, AC004815, AL031257, AC007263, and Z83844.
HYAAB40	6876	841969	1 - 631	15 - 645	AA497098, and AC006023.
HWWEW11	6877	965434	1 - 648	15 - 662	W38418, W38417, AW362370, AI368729, AI523296, AA548224, AA868898, AI653483, W73847, AA558195, AA700476, AI828445, AW305153, AA548241, AA707354, W69719, W73703, AA164647, AI539088, Z19840, AI752812, AI752779, AW262145, W38359, AI399680, AA164646, R71097, AI073454, AI871525, F09131, AI686475, AA917430, AA372471, AI811402, R71445, AA418958, AI695645, AI242956, AA385654, AI304631, AI597928, AW362391, and A75407.
HWWEL11	6878	965426	1 - 423	15 - 437	AA553392, AA526829, AI023375, AC008372, AC005839, AF111168, AC005288, AC007371, AC004520, AC000353, AL031733, AL022163, AL022320, AC005756, AC004983, AC004707, AL121652, AC002300, AL109984, AC006581, AC004966, AJ003147, AC005510, and AL031276.
HWWEC06	6879	933690	1 - 415	15 - 429	AW452648, AI249134, AA831307, AA504653, AW182255, AI468242, AW444889, H14382, T33198, AA731107, F37077, R99416, W46344, W05246, AI632386, AI521519, AW150858, and AF151907.
HWBGG82	6880	779081	1 - 665	15 - 679	N23897, and H84937.

What Is Claimed Is:

1. An isolated nucleic acid molecule comprising a polynucleotide having a nucleotide sequence at least 95% identical to a sequence selected from the group consisting of:

(a) a polynucleotide fragment of SEQ ID NO:X or a polynucleotide fragment of the cDNA sequence contained in Clone ID NO:Z, which is hybridizable to SEQ ID NO:X;

(b) a polynucleotide encoding a polypeptide fragment of SEQ ID NO:Y or a polypeptide fragment encoded by the cDNA sequence contained in cDNA Clone ID NO:Z, which is hybridizable to SEQ ID NO:X;

(c) a polynucleotide encoding a polypeptide fragment of a polypeptide encoded by SEQ ID NO:X or a polypeptide fragment encoded by the cDNA sequence contained in cDNA Clone ID NO:Z, which is hybridizable to SEQ ID NO:X;

(d) a polynucleotide encoding a polypeptide domain of SEQ ID NO:Y or a polypeptide domain encoded by the cDNA sequence contained in cDNA Clone ID NO:Z, which is hybridizable to SEQ ID NO:X;

(e) a polynucleotide encoding a polypeptide epitope of SEQ ID NO:Y or a polypeptide epitope encoded by the cDNA sequence contained in cDNA Clone ID NO:Z, which is hybridizable to SEQ ID NO:X;

(f) a polynucleotide encoding a polypeptide of SEQ ID NO:Y or the cDNA sequence contained in cDNA Clone ID NO:Z, which is hybridizable to SEQ ID NO:X, having biological activity;

(g) a polynucleotide which is a variant of SEQ ID NO:X;

(h) a polynucleotide which is an allelic variant of SEQ ID NO:X;

(i) a polynucleotide which encodes a species homologue of the SEQ ID NO:Y;

(j) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i), wherein said polynucleotide does not hybridize under stringent conditions to a nucleic acid molecule having a nucleotide sequence of only A residues or of only T residues.

2. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises a nucleotide sequence encoding a protein.

3. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises a nucleotide sequence encoding the sequence identified as SEQ ID NO:Y or the polypeptide encoded by the cDNA sequence contained in cDNA Clone ID NO:Z, which is hybridizable to SEQ ID NO:X.

4. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises the entire nucleotide sequence of SEQ ID NO:X or the cDNA sequence contained in cDNA Clone ID NO:Z, which is hybridizable to SEQ ID NO:X.

5. The isolated nucleic acid molecule of claim 2, wherein the nucleotide sequence comprises sequential nucleotide deletions from either the C-terminus or the N-terminus.

6. The isolated nucleic acid molecule of claim 3, wherein the nucleotide sequence comprises sequential nucleotide deletions from either the C-terminus or the N-terminus.

7. A recombinant vector comprising the isolated nucleic acid molecule of claim 1.

8. A method of making a recombinant host cell comprising the isolated nucleic acid molecule of claim 1.

9. A recombinant host cell produced by the method of claim 8.

10. The recombinant host cell of claim 9 comprising vector sequences.

11. An isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence selected from the group consisting of:

(a) a polypeptide fragment of SEQ ID NO:Y or the encoded sequence contained in cDNA Clone ID NO:Z;

(b) a polypeptide fragment of SEQ ID NO:Y or the encoded sequence contained in cDNA Clone ID NO:Z, having biological activity;

(c) a polypeptide domain of SEQ ID NO:Y or the encoded sequence contained in cDNA Clone ID NO:Z;

(d) a polypeptide epitope of SEQ ID NO:Y or the encoded sequence contained in cDNA Clone ID NO:Z;

(e) a full length protein of SEQ ID NO:Y or the encoded sequence contained in cDNA Clone ID NO:Z;

(f) a variant of SEQ ID NO:Y;

(g) an allelic variant of SEQ ID NO:Y; or

(h) a species homologue of the SEQ ID NO:Y.

12. The isolated polypeptide of claim 11, wherein the full length protein comprises sequential amino acid deletions from either the C-terminus or the N-terminus.

13. An isolated antibody that binds specifically to the isolated polypeptide of claim 11.

14. A recombinant host cell that expresses the isolated polypeptide of claim 11.

15. A method of making an isolated polypeptide comprising:

(a) culturing the recombinant host cell of claim 14 under conditions such that said polypeptide is expressed; and

(b) recovering said polypeptide.

16. The polypeptide produced by claim 15.

17. A method for preventing, treating, or ameliorating a medical condition, comprising administering to a mammalian subject a therapeutically effective amount of the polynucleotide of claim 1.

18. A method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject comprising:

(a) determining the presence or absence of a mutation in the polynucleotide of claim 1; and

(b) diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or absence of said mutation.

19. A method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject comprising:

(a) determining the presence or amount of expression of the polypeptide of claim 11 in a biological sample; and

(b) diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or amount of expression of the polypeptide.

20. A method for identifying a binding partner to the polypeptide of claim 11 comprising:

(a) contacting the polypeptide of claim 11 with a binding partner; and

(b) determining whether the binding partner effects an activity of the polypeptide.

21. The gene corresponding to the cDNA sequence of SEQ ID NO:Y.

22. A method of identifying an activity in a biological assay, wherein the method comprises:

(a) expressing SEQ ID NO:X in a cell;

(b) isolating the supernatant;

(c) detecting an activity in a biological assay; and

(d) identifying the protein in the supernatant having the activity.

23. The product produced by the method of claim 20.

24. A method for preventing, treating, or ameliorating a medical condition, comprising administering to a mammalian subject a therapeutically effective amount of the polypeptide of claim 11.